Local-global processing and cognitive style in autism spectrum disorders and typical development

Booth, Rhonda Denise Lowsley

LOCAL-GLOBAL PROCESSING AND COGNITIVE STYLE IN AUTISM SPECTRUM DISORDER AND TYPICAL DEVELOPMENT

R D L Booth

Institute of Psychiatry

King's College London

Thesis submitted for the degree of
Doctor of Philosophy

University of London

2006

ABSTRACT

This thesis was motivated by the hypothesis that a continuum of cognitive style may exist in the general population, from strong to weak coherence. On this conceptualisation, individuals with autism spectrum disorder (ASD) are proposed to lie at the extreme weak coherence or detail-focused end of the continuum. A battery of 14 coherence measures was administered to over 200 typically developing (TD) individuals, 31 individuals with ASD, and 31 age- and IQ-matched control participants. The pervasiveness of central coherence was tested across visuo-spatial and auditory/verbal domains, and high- and low-levels of processing.

Age and IQ were typically related to task performance in the TD group, but a large proportion of variance remained unexplained and may reflect cognitive style. Low-level tasks were associated within and across visual and auditory domains, suggesting some consistency in individual differences. High-level tasks did not show such consistency, suggesting that executive/strategic processes may have greater effect on task performance than local-global processing style. Males showed greater detail-focus and stronger developmental effects (more global with age) than females on several measures.

Weak coherence was demonstrated in ASD by local processing bias and lack of global bias. Local and good global processing appeared to be more in trade-off in the ASD group than in TD and control groups. Good local processing related to IQ in TD and controls, but less so in ASD suggesting local bias is more automatic to individuals with ASD.

Subgroups were determined on the basis of performance across the battery, according to whether local or global processing was dominant, or whether an individual adapted well or poorly to the demands of the task. A consistent local processing style was more common in the ASD group than in the control group, but was not universal. The implications of these findings for weak central coherence theory of ASD are discussed.

ACKNOWLEDGMENTS

I am extremely privileged to have Dr Francesca Happé as a mentor and I would like to give my sincere thanks to her unfailing support and encouragement throughout the PhD process and beyond.

I am forever grateful to the teachers who showed an interest in the study and gave me permission to move into their school for months at a time, and provided invaluable help with recruitment. Special thanks to Angela Daly and Brigitte Santer at Cranbrook, Stephanie Bedford and Lesley Sendles at Angley, O'Neil Hemming and Alison Honeysett at St Saviour's Primary, Darra McFayden at Jessop Primary, Rosemary Newlove and Debbie Strode at Streatham Wells Primary, Peter Blewett at Warwick Park, Andy Cobley and Adrian Wylie at North Hill House, and David Cassar and Deirdre Goodger at St Mary's. Thanks also to Dr Fred Seymour, Psychology Department, University of Auckland.

I am extremely thankful to all my participants who were so willing to take part and allow me insight into their minds.

I would like to thank Drs Morven Leese and Frühling Rijsdijk for their statistical advice; Dr Pamela Heaton for her expertise in music cognition; John Rogers for his programming skill and brainstorming sessions on task design; Greg Wallace and Eva Loth for theoretical discussions, and Antonia San José Cáceres and Rebecca Charlton for their many hours of double-coding.

My final thanks goes to my family, friends and colleagues who gave me much support throughout (Alessandra, Essi, Kathryn, Kerri, Mette, Steve) and especially to Illtyd Gwyn Harri Griffiths for his enduring faith and for being there.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction to Autism Spectrum Disorder

1.1 Introduction

This thesis explores a concept first introduced in the autism literature; that individuals differ in their drive for central coherence, an information-processing bias that varies from 'local' detail-focused to strong 'global' coherence. A brief overview of autism spectrum disorder (ASD) is therefore necessary before moving onto a more detailed review of central coherence theory and related concepts. This first chapter provides an overview of what is known about ASD today and introduces the reader to current knowledge regarding the prevalence, neurobiological and neurophysiological underpinnings, and cognitive theories of this disorder.

1.2 Definition and diagnosis

Autism is a profound and complex developmental disorder with the essential feature being marked impairments in the way an individual communicates and interacts with other people. The term *early infantile autism* was first used over 60 years ago by the American psychiatrist Leo Kanner to describe a group of referred children who all had "extreme aloneness from the very beginning of life" (Kanner, 1943, p. 248). Kanner collated the puzzling combination of symptoms presented in these children, which included impaired social responsiveness, an obsessive desire for the "preservation of sameness", echolalia, poor eye contact, restricted interests, and oversensitivity to stimuli, combined with good memory and seemingly good cognitive potential. The expression *autistic psychopathy* was independently used by the Austrian paediatrician Hans Asperger (1944, translated by Frith, 1991) to describe four children showing a similar pattern of social withdrawal and obsessive interests. There were differences between the two descriptions, most notably Asperger observed fluent language abilities, poor motor coordination, and evidence of abstract thought (Wing, 1991).

Kanner and Asperger's early descriptions formed the conceptualisation of autism as it is known today, although modifications have been made as more has been learnt about the disorder. It is recognised that autism can have a range of behavioural manifestations, from mild to severe, with or without additional handicaps and mental retardation. Wing and Gould (1979) introduced the concept of an autistic spectrum, to cover a range of ability levels and severities, all characterised by qualitative impairments in social, communication and imaginative development. These three domains of impairment (commonly referred to as the 'triad of impairments') form the basis of the current diagnostic criteria in the *Diagnostic and Statistical Manual of the Mental Disorders, 4th edition* (DSM IV; American

Psychiatric Association, 1994) and the *International Classification of Diseases, 10th revision* (ICD 10; World Health Organization, 1993). Both diagnostic systems agree that qualitative impairments in development must be present prior to the age of three. The core symptoms for diagnosis include: (1) qualitative impairment in social interaction as shown by impaired use of non-verbal behaviours such as eye-gaze and joint attention, failure to develop peer relationships, and lack of social/emotional reciprocity; (2) qualitative impairments in communication as shown by a delay in, or lack of, development of spoken language or language characterised by idiosyncrasies such as pronoun reversal and echolalia, marked impairment in the ability to initiate or sustain a conversation; and (3) stereotyped patterns of behaviour, interests and activities, such as repetitive motor mannerisms, insistence on sameness and inflexibility.

For a diagnosis of Autistic Disorder impairments in all three domains must be present. Asperger's Disorder, although controversially seen as a distinct disorder (see review by Frith, 2004), is included on the autism spectrum and is diagnosed on the basis of impairments in communication, social behaviour and rigid behaviour, but with no clinically significant general delay in language. A diagnosis of Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) is given when a severe and pervasive impairment in the development of social interaction or communicative skills, or when stereotyped patterns of behaviour and repetitive interests are present, but not all the criteria for Autistic Disorder or Asperger's Disorder diagnosis are met. All three disorders are classified under the broader heading of Pervasive Developmental Disorders (PDD), which acknowledges autism as a lifelong disorder with persistent behavioural symptoms. ASD is the collective term used to include the range of manifestations of the disorder, and will be used throughout this thesis to refer to individuals who have a diagnosis of Autistic Disorder, high-functioning autism (HFA) or Asperger's Disorder.

1.3 Prevalence

In a recent review of epidemiological surveys, Rutter (2005) surmised the current rates of ASD are likely to be in the region of 30-60 per 10,000 with approximately a quarter of those cases meeting the full criteria for autism. This is a marked increase on estimates provided by earlier studies (e.g., 4 per 10,000; Lotter, 1966) and has evoked claims of an 'epidemic' of autism (see review and discussion by Fombonne, 2003, 2005). Prevalence estimates are likely to depend on assessment tools and ascertainment methods, with variations across studies reflecting differences in methodology. Furthermore, Rutter maintains the increase in prevalence and incidence is primarily due to the broadening of the diagnostic concept in addition to changes in diagnostic practice and public and professional awareness.

Autism affects more males than females (e.g., 4 to 1; Chakrabarti & Fombonne, 2001) and this discrepancy becomes more pronounced at the more able end of the spectrum (e.g., 15 to 1; Wing, 1981). Although autism can occur at all levels of intellectual ability, approximately 75% of people with autism are reported to be low-functioning (Rutter, 1978). This percentage may be much lower in more recent samples of the full autism spectrum (i.e., 25%; Baird et al., 2000).

1.4 Aetiology

Several clinical features provide evidence that ASD has a neurobiological basis: the higher than expected incidence of seizures (ranging from 11 to 42%; Olsson, Steffenburg, & Gillberg, 1988; Volkmar & Nelson, 1990); abnormalities on electroencephalograms (ranging from 15% to 36% of children with ASD and without epilepsy; Tuchman & Rapin, 1997); the association with mental retardation, and the sex ratio in favour of males (mentioned above). The exact aetiology of autism is currently unknown, although there is consensus of a large genetic component in the majority of cases. A variety of causes are possible, involving complex interaction among genes and environment.

Environmental factors such as infections, toxins, or nutrition, cannot be entirely excluded as a causal factor to ASD but there is little research evidence to date in support of such risk factors. The possible causal effects of the triple measles, mumps and rubella (MMR) vaccine have largely been discounted (Honda, Shimizu, & Rutter, 2005). Perinatal complications are considered more likely to be consequences than causes of ASD (Bolton, Murphy, Macdonald, & Whitlock, 1997; Zwaigenbaum et al., 2002). No specific prenatal exposures have been recognised as contributory (see Muhle, Trentacoste, & Rapin, 2004 for a review), with the possible exception of thalidomide (Stromland, Nordin, Miller, Akerstrom, & Gilberg, 1994) and rubella early in gestation (Chess, Fernandez, & Korn, 1978). Postnatal herpes encephalitis has also been described as associated with autism (Gillberg & Coleman, 1992).

1.4.1 Genetic factors

Autism is considered to be one of the most heritable developmental disorders, with estimates of heritability exceeding 90% (Bailey et al., 1995). Relatives of individuals with ASD are more likely to exhibit behaviours and cognitive biases notable in autism, albeit of a lesser degree, providing evidence for the existence of a 'broader phenotype' for autism (Bailey, Palferman, Heavey, & Le Couteur, 1998; Dawson et al., 2002).

Support for a genetic component in ASD also comes from association with several disorders that have established genetic causes such as tuberous sclerosis (Baker, Piven, & Sato, 1998; Smalley, 1998) and fragile X syndrome (Bolton & Rutter, 1990). Although these syndromes account for a small proportion of all cases of autism, they may provide important information about the genetic cause of the disorder. The search for relevant genes is ongoing, but it is considered that multiple genes are involved (see Maestrini, Paul, Monaco, & Bailey, 2000, for reviews of susceptibility genes), that interact with several environmental factors (Pickles et al., 1995). The heterogeneity of ASD also poses problems in establishing the genetic loci and linkage findings have been strengthened by examination of distinct subgroups, for example, those with language impairments or savant skills (Bradford, 2001; Nurmi et al., 2003; Shao et al., 2002).

1.4.2 Neuroanatomic basis

A selective review of neurobiological findings in ASD at anatomical and physiological levels is presented (for more comprehensive reviews see Bauman & Kemper, 2005; Volkmar, Lord, Bailey, Schultz, & Klin, 2004). Although few post-mortem studies are available, an increasing number of magnetic resonance imaging (MRI) studies have been conducted, and together report structural brain abnormalities in autism. In a recent review, Bauman and Kemper (2005) reported that consistent findings have been observed in areas of the limbic system, the cerebellum and related inferior olive. In a review of studies published between 1966 and 2003, Brambilla et al. (2003) concluded that increased total brain, parieto-temporal lobe, and cerebellar hemisphere volumes were the most replicated abnormalities reported. The amygdala, hippocampus, and corpus callosum have also shown irregularities in structure.

Recent emphasis seems to be turning towards abnormalities of connectivity, rather than discrete areas of damage (see Belmonte et al., 2004, for a review). The increase in overall brain size in autism in early childhood compared to normal controls is an intriguing finding and has been recently confirmed through a meta-analysis of all studies reporting brain size and volume (Redcay & Courchesne, 2005). The implications of this finding and the underlying mechanisms involved have not been ascertained, although there is speculation that neuronal pruning in autism is impaired (Courchesne et al., 2001; C. Frith, 2004). Although not conclusive, the anatomical evidence gathered to date suggests that abnormal brain development in autism occurs early in development (Bauman & Kemper, 2005).

1.4.3 Neurophysiological abnormalities

An increasing number of brain imaging studies in autism using functional MRI and position emission tomography (PET) techniques have appeared in recent years (for a review see Cody, Pelphrey, & Piven, 2002). Studies have reported distinct functional abnormalities in a number of cortical (notably frontal and temporal lobes and the cerebellum) and subcortical regions (notably amygdala and hippocampus) as individuals with ASD performed psychological tasks, although the findings are somewhat inconsistent (e.g., Baron-Cohen et al., 1999; Hadjikhani et al., 2004; Happé et al., 1996; Haznedar et al., 2000; Muller et al., 1999; Muller, Pierce, Ambrose, Allen, & Courchesne, 2001; Ring et al., 1999).

1.5 Cognitive theories of ASD

The puzzling cluster of symptoms in ASD has motivated a large number of theories concerning the underlying cognitive impairment (see Happé, 1994a; Hill & Frith, 2003, for reviews). Theories broadly fall into two types: those that suggest the primary impairment is in social development, and those that propose non-social deficits are fundamental and have a secondary impact on social adaptation. Two main theories emanating from social (theory of mind) and non-social (executive function) domains will be outlined in brief. Although influential, both theories are deficit accounts and cannot fully explain the full range of behaviours exhibited by individuals with ASD. This paves the way to a third theory (weak central coherence), which attempts to explain both assets and deficits of ASD and, as the main focus of this thesis, will be discussed in detail in Chapter 2.

1.5.1 The Theory of Mind hypothesis

The most influential theory in recent years is the suggestion that individuals with ASD are impaired in 'theory of mind' or 'mentalising,' that is, the ability to attribute independent mental states to self and others in order to explain and predict behaviour (see Frith, 2003, for review). Individuals with autism are hypothesised to be unable to represent the intentions, desires, feelings and beliefs of other people and, as a consequence, have difficulties in social interaction and communication. This theory has been successful in explaining the core socio-communication and imaginative impairments seen in ASD, and the division between intact and impaired areas of development (see Baron-Cohen, Tager-Flusberg, & Cohen, 2000, for reviews).

The theory of mind hypothesis has undergone rigorous examination (Baron-Cohen et al., 2000) and several limitations to the explanatory power of the theory have been made. One limitation is that a significant minority of individuals with ASD pass standard theory of mind tasks (e.g., false belief tasks) but still present severe social impairments. Thus severity of ASD symptoms cannot be predicted by experimental test performance. Success on theory of mind tasks has been found to relate strongly to verbal ability in ASD (Happé, 1995). This suggests that 'mind-reading' capabilities in ASD are not as automatic or implicit as in typical development and may be acquired at a later stage of development or are accomplished through alternative routes (e.g., logical reasoning). The primacy of theory of mind impairments has also been questioned. Deficits in theory of mind may therefore be an adequate description of the end point of abnormal social development, and other factors, such as a reduced tendency to orient to social cues (Klin, Jones, Schultz, Volkmar, & Cohen, 2002) may be more fundamental.

The theory of mind account of autism has had great influence in directing both research and clinical practice; for example in investigating the neurophysiological substrates of mentalising (see Frith, 2001, for a review), developing tools for early detection such as the Checklist for Autism in Toddlers (CHAT; Baird et al., 2000), and in intervention studies which have shown some success in the remediation of theory of mind abilities, but with limited generalisation (Howlin, Baron-Cohen, & Hadwin, 1998; Swettenham, Baron-Cohen, Gomez, & Walsh, 1996). Performance on theory of mind tasks has also been used to characterise the broader phenotype of autism (Klin et al., 2002). Despite this impact, there are aspects of ASD which a deficit in theory of mind cannot account for, in particular repetitive behaviour, the uneven cognitive profile and 'islets of ability' often seen in people with autism.

1.5.2 Executive function deficits

In addition to the triad of impairments in socialisation, communication and imagination, the diagnostic criteria of ASD include restricted, repetitive and stereotyped patterns of behaviour. Parallels have been drawn between these inflexible and perseverative behaviours and the symptoms of patients with frontal lobe injury, leading to studies of executive functioning in autism (Hill, 2004). 'Executive function' is an umbrella term for the set of higher-order cognitive processes required for adaptive responses to novel situations. These include set-shifting, inhibition of prepotent responses, self-monitoring and planning, and are generally impaired by damage to the prefrontal cortex (Ozonoff et al., 2004; Pennington & Ozonoff, 1996). In a comprehensive review, Hill (2004) concluded that impairments in ASD (compared to typically developing controls) are often found on tasks of planning, mental flexibility and generativity, while the evidence for inhibition and self-monitoring deficits was mixed. Executive functioning problems have been found in biological relatives of individuals with autism (e.g., Hughes, Leboyer, & Bouvard, 1997; Ozonoff, Rogers, Farnham, & Pennington, 1993), and suggest the broader phenotype of autism may include deficits in the executive domain.

The major failing of the theory is that executive function deficits are not specific to ASD and are common in other developmental disorders, such as Attention Deficit Hyperactivity Disorder (ADHD), conduct disorder, and Tourette syndrome. This limits the possibility of using executive dysfunction as a diagnostic marker for autism or as an explanation for the characteristic (social) symptom profile. Studies are now including clinical groups with known executive dysfunction to compare performance with ASD (e.g., Geurts, Verte, Oosterlaan, Roeyers, & Sergeant, 2004; Happé, Booth, Charlton, & Hughes, 2006) which may determine whether specific executive performance profiles can differentiate clinical groups.

A further problem is that while impairments in executive functioning are common in ASD, they are not a universal feature (Hill, 2004). In particular, executive functioning problems may not differentiate preschoolers with ASD from ability-matched controls with developmental delay (Griffith, Pennington, Wehner, & Rogers, 1999). Despite the issues of specificity and universality, research into executive functioning has helped define the difficulties individuals with ASD encounter with everyday planning, decision-making and adapting to change, and has started to direct the course of intervention studies (Fisher, 2002; Shimmon & Lewis, 2001).

1.6 Summary

Since Kanner's first description of the syndrome of early infantile autism, the field of research into ASD has grown exponentially. While early descriptors have proved remarkably robust, these continue to be refined as more is known about the disorder. Investigations into the neurophysiological basis of ASD have also increased in recent years and it is hoped that future diagnostic techniques will be able to include genetic, neuroimaging, and neurochemical markers. Research into ASD is often impeded by the high heterogeneity of the disorder and further study into the cognitive, behavioural and biological characteristics may help determine whether we are studying a unitary group of individuals.

Challenges still remain for cognitive theories of ASD. In particular, several features of autism cannot easily be explained by current deficit accounts (theory of mind, executive dysfunction). These include restricted repertoire of interests, obsessive desire for sameness, islets of ability, excellent rote memory and preoccupation with parts of objects. The proposal made by Frith (1989), suggesting individuals with autism have a weak drive for coherence, was offered to explain the specific pattern of deficits and abilities associated with this disorder. Chapter 2 presents a review of the current status of this theory, documenting empirical evidence for and against weak central coherence as a possible cognitive marker for ASD.

Chapter 2. The weak central coherence account of autism

2.1 Introduction

It has been proposed that individuals with autism can be characterised by a specific processing style, that is, a bias towards local detail in preference to global form. Uta Frith introduced the term "weak central coherence" to describe this unusual cognitive style in her seminal book *Autism: Explaining the Enigma* published in 1989. Since its inception, a wealth of research has emerged that has examined whether weak central coherence is a valid and reliable theory. It has been questioned whether weak central coherence is specific and universal to individuals with ASD, whether it can explain, or be explained by, other established theories of autism, whether alternative accounts are more plausible, and what the underlying mechanisms may be. The following chapter provides a critical review of empirical studies that have tested the weak central coherence theory of autism. Particular attention is given to studies that have addressed the two main questions of this thesis: (1) Is weak coherence universal and specific to ASD? (2) Do individual differences exist in coherence and are they pervasive across different processing levels and domains?

2.2 Clinical impressions of fragmented perception in autism

Clinical descriptions and autobiographical accounts provide rich descriptions of the unusual sensory and perceptual experiences of individuals with ASD. Observations of children with autism often report the tendency to notice minute details and to react to very small changes in their environment (Wing, 1976). Frith and Baron-Cohen (1987) described the unusual responsiveness to stimuli in a boy with autism who "could always find small coins long before anyone else in the room had seen them, although he appeared oblivious to many other visual aspects of the environment" (pp. 86 – 87). This unusual attention to detail was highlighted in Kanner's (1943) original description of autism: "a situation, a performance, a sentence is not regarded as complete if it is not made up of exactly the same elements that were present at the time the child was first confronted with it" (p. 246). The characteristic 'insistence on sameness', described by Kanner, such that the sight of anything broken or incomplete could lead to great distress in these children, may be seen as the result of this excessive focus on parts rather than the whole. The observation of "persistent preoccupation with parts of objects" is one of the current diagnostic criteria for autistic disorder (DSM-IV; American Psychiatric Association, 1994). This may manifest in an unusual tendency to play with specific parts of toys (e.g., spinning the wheels of a car), rather than according to its intended function (Lord, Rutter, & Le Couteur, 1994).

Further examples of perceptual peculiarities in autism have come from autobiographical writings (e.g., Gerland, 1997; Grandin, 2000; D. Williams, 1992). Intense sensitivity to sound, touch or vision are often described, as well as an acute experience of details in the environment that typically go unnoticed. In one personal account, Williams described her senses and perception as "chaotic, fragmented and constantly shifting and fluctuating." Similarly, Gerland described a fragmented perception of the world. She considered her visual perception to be two-dimensional and suffered problems with depth perception and perspective. Gerland also encountered difficulties with high-level verbal integration, as described when attempting to piece together facts to write an essay: "Every little bit of fact seemed to land in its own compartment in my head and refused to be linked with any other. I tried poking into details. I dissected them and hoped a unified whole would appear, but it rarely did" (p. 225).

2.3 The weak central coherence proposal

The perceptual and cognitive abnormalities seen in people with autism were proposed to stem from a disruption in normal information processing. Frith (1989) coined the phrase "central coherence" to described the natural tendency of the mind to pull information together. A drive for coherence enables one to make sense out of information, build up context and see structure and meaning, although to the detriment of detail and surface form. Frith suggested that this capacity for central coherence was diminished in individuals with autism who instead show "weak central coherence", a drive towards local detail, at the expense of global meaning.

Rather than being a deficit account, weak central coherence predicts relatively good performance where attention to local information and ignoring the context is advantageous, and conversely, poor performance on tasks that require integration of information in context or recognition of global meaning. As weak central coherence may facilitate certain strengths in autism, such as special talents and savant skills (Hermelin, 2001), Happé (1999) proposed that the notion of a core deficit in coherence should be replaced by a processing *style*. Indeed, it appears this style or bias can be overcome in tasks that explicitly demand global processing: individuals with autism can process information for meaning, when instructed to do so, but it may not be their natural tendency. As a cognitive style, Happé proposed that this propensity for local versus global processing might vary along a continuum in the normal population, varying from "strong" to "weak" coherence. Individuals with autism are hypothesised to be at the extreme (weak) end of this distribution. Weak coherence is postulated to be one among several cognitive anomalies in ASD, with social deficits, for example, explained by distinct and independent theory of mind impairments.

2.4 Empirical evidence of weak central coherence in ASD

A number of empirical studies have shown that individuals with ASD, compared to control participants, show a reduced tendency to cohere information presented at basic perceptual levels in visual and auditory domains, as well at higher semantic levels during verbal and visuo-spatial problem-solving. Central coherence has been assessed in various ways: comparing performance on meaningful versus meaningless stimuli, comparing global versus local processing, and comparing preference for gestalt versus fragmented patterns. A table of published experimental studies that have addressed weak central coherence in ASD is provided by Happé and Frith (2006). A selection of major studies from the domains of low-level perceptual, visuo-spatial, verbal-semantic, and music processing, are reviewed below.

2.4.1 Perceptual coherence

The perception of visual illusions has been used to explore weak central coherence in autism at a low level of perceptual processing. Happé (1996) reported that children with autism did not readily succumb to illusory effects and made more accurate judgments of the physical properties of stimuli than typically developing (TD) children and those with moderate learning difficulties (MLD). This finding was suggested to arise from a failure to integrate the parts of the figure that would typically induce an illusion. This was confirmed by a control condition where the same illusions were presented in an enhanced, pre-segmented format (three-dimensional). Control participants became more accurate in their judgments in this condition, but the pre-segmentation did not aid the individuals with autism. This resistance to visual illusions has not been replicated, however, in studies by Ropar and Mitchell (1999; 2001) and more recently by Hoy, Hatton and Hare (2004). Using a more sophisticated measure whereby the participant could adjust the physical dimensions of an illusion until they appeared perceptually equivalent, Ropar and Mitchell found individuals with autism performed in the same manner as controls. Brosnan, Scott, Fox, and Pye (2004) comment that the nature of the verbal instructions given can impact on whether individuals with ASD succumb to visual illusions. They found individuals with autism tended to succumb to the Muller–Lyer illusion when asked, "which line looks longer" but did not succumb when asked "which line is longer."

Although a lack of susceptibility to visual illusions may not be a robust finding, other studies on coherence at the perceptual level have provided evidence for a local processing bias in individuals with autism. Gepner, Mestre, Masson, and de Schonen (1995) demonstrated that five children with autism were less susceptible to visually induced motion than age-matched TD children. They later replicated this finding in autism (although not in Asperger syndrome; Gepner & Mestre, 2002), which suggests that individuals with autism are indifferent to global changes in the environment that typically induce postural instability.

Also at a perceptual level of coherence, Jarrold and Russell (1997) compared dot-counting abilities between children with autism, children with MLD, and TD children. Children with autism were much slower to count dots when they appeared in canonical patterns (as on dice), than when the dots were distributed in a non-uniform pattern. They tended to count each dot individually and not benefit from the global pattern of presentation, suggesting a tendency towards an analytic level of processing. Similarly, Brosnan et al. (2004) found children with autism used gestalt grouping principles of proximity, similarity, and closure less often than children with learning difficulties. They concluded that children with autism were grouping at a level that was not significantly greater than chance, resulting from a difficulty in perceiving relationships between component parts.

2.4.2 *Visuo-spatial coherence*

An early study providing the impetus to the weak central coherence theory was reported by Shah and Frith (1983). Individuals with autism demonstrated less capture by meaning and gestalt on the Children's Embedded Figures Task (CEFT), which requires the identification of simple figures embedded within a complex form. Individuals with autism could quickly and accurately locate the embedded figures, compared to control participants, and this was interpreted as an ability to ignore the strong gestalt of the complex form, allowing effortless perception of local parts. Superior performance on the Block Design task (in which a pattern has to be constructed from individual blocks) has been demonstrated in ASD and this again, has been attributed to a local, rather than global, processing style. To confirm this, Shah and Frith (1993) found the relative benefit of presenting designs in a segmented form compared to the whole form was reduced in ASD compared to normal and mental-age (MA) matched controls. They proposed that this provided evidence of weak central coherence in autism; these individuals readily perceived the designs in terms of their constituent parts and were not locked into the strong 'gestalt' of the design.

Individuals with autism have been found to perform well when required to recognise objects from individual fragments, but encounter difficulties when required to integrate the fragments into a whole (Jolliffe & Baron-Cohen, 2001a). The ability to visually integrate local parts into a global form is necessary to determine whether a three-dimensional figure is geometrically possible or impossible (Young & Deregowski, 1981). Individuals with autism have been found to encounter difficulties on such tasks (Rodgers, 2000), and when asked to copy an impossible figure in a drawing task, were not hindered by their globally incoherent form (Mottron & Belleville, 1993).

Unusual, detail-focused drawing styles have been observed in individuals with ASD. Mottron and colleagues reported the tendency to begin drawings with a local feature, rather than sketch a global outline, in adolescents and adults with autism (Mottron & Belleville, 1993; Mottron, Belleville, & Menard, 1999). Booth, Charlton, Hughes and Happé (2003) identified several markers for weak central coherence in a simple drawing task that were more apparent in ASD than in an age- and ability-matched control group. Boys with ASD were more likely to begin with local elements or details, draw in a piecemeal or fragmented fashion, and violate the overall configuration of the figure. Fein, Lucci and Waterhouse (1990) also explored fragmentation in drawing, as well as overlap of drawn parts. They found more evidence of these signs of failure to integrate the whole in an autism group (aged 5 to 17 years) compared to MA-matched control children when asked to draw a person.

Studies utilising Navon hierarchical figures (e.g., an H composed of small Ss; Navon, 1977) have produced mixed results. The typical global precedence effect, that is, a processing advantage for global level stimuli over local level stimuli, has been found in children with ASD (Ozonoff, Strayer, McMahon, & Filloux, 1994) contrary to expectations from weak central coherence theory, which would predict at least reduced global precedence. However, global precedence effects in individuals with typical development have not been reliably demonstrated in these studies (e.g., Mottron, Burack, Stauder, & Robaey, 1999). Navon tasks are known to be sensitive to small variations in methodology (Kimchi, 1992), which may explain conflicting findings. In Ozonoff et al.'s study for example, attention was directed specifically to the local and global levels in this task, which may not tap the natural processing preference of individuals with ASD. To resolve this issue, Plaisted, Swettenham, and Rees (1999) examined both divided and selective attention paradigms, and found that when attention was divided across local and global levels, individuals with ASD showed a local advantage and interference from local to global stimuli. However, Rinehart, Bradshaw, Moss, Brereton, and Tonge (2000) were able to demonstrate local interference in individuals with ASD using a selective attention design.

2.4.3 Verbal-semantic coherence

Evidence of local processing bias was observed in early studies of verbal memory and language processing in autism. In the pioneering work of Hermelin and O'Connor (1970), children with autism were found not to derive the typical benefit from meaning in memory tests. In particular, recall for words did not improve as much as for control children when the words were arranged into sentences compared to random word lists. Furthermore, children with autism tended not to reorder words into semantic clusters in order to facilitate recall. This finding has since been replicated (e.g., Tager-Flusberg, 1991) and suggests that individuals with autism do not make use of the semantic organisation of verbal information to aid memory.

Individuals with autism have also been shown to be impaired in the use of context when reading homographs presented in sentences (e.g., "In her dress/eye there was a big tear"; Frith & Snowling, 1983; Happé, 1997; Jolliffe & Baron-Cohen, 1999). It has been shown that people with autism do not use context to guide their pronunciation of homographs, often giving the more frequent pronunciation irrespective of the contextually-determined meaning. A similar disregard for sentence context has been demonstrated on the Sentence Completion Task (Happé & Booth, 2006), where individuals with ASD at all levels of ability, show a greater tendency to make local, globally inappropriate, completions to sentence stems than control participants. For example, when asked to complete "*You can go hunting with a knife and ....*" a local response such as "*fork*" would be given, rather than a globally meaningful response such as "*catch a bear*".

The lack of drive for meaning has also been shown at the narrative level. Individuals with autism have demonstrated good verbatim but poor gist memory for story material (Scheuffgen, 1998). Adults with autism were less able to arrange sentences coherently and use context to make a global inference than TD controls (Jolliffe & Baron-Cohen, 2000). The reduced influence of context may have benefits however, with individuals with autism demonstrating less susceptibility to false memories; that is, memories consistent with the interpreted gist of a situation but not the precise details. Weak coherence may lead to greater accuracy in recall in ASD, possibly as a result of a reduced influence of context (Beversdorf et al., 2000; although see Bowler, Gardiner, Grice, & Saavalainen, 2000) .

2.4.4 Music processing

Investigations into local and global processing of musical stimuli in individuals with ASD have shown inconsistent findings (Foxton et al., 2003; Heaton, 2005; Heaton, Pring, & Hermelin, 1999; Mottron, Peretz, & Menard, 2000). In a commonly used paradigm, participants are asked to make same/different judgments about pairs of melodies. Comparison melodies are made different by altering the pitch of one tone, which either violates the contour (i.e., a *global* modification) or maintains the contour of the melody (i.e., a *local* modification). Previous studies in typical populations have shown that it is difficult to detect differences between melody pairs that maintain global characteristics such as overall contour and harmonic structure (Trehub, Schellenberg, & Hill, 1997). It was hypothesised that individuals with autism, with a bias towards focusing on individual elements, may detect the subtle local changes even when the global characteristics are preserved.

Heaton et al. (1999, Experiment 5) did not confirm this prediction, however, and found that a group of children with ASD, like their age- and IQ-matched controls, processed the melodies in terms of their global form. Mottron et al. (2000) used this same paradigm with a group of high-functioning adolescents and adults with autism and age- and nonverbal IQ-matched controls. Both groups found global modifications easier to detect than local modifications. Individuals with autism were generally better than control participants at detecting differences, especially for melodies with local modifications. The authors concluded that individuals with autism demonstrated intact global processing alongside a local bias in music processing. Foxton et al. (2003) included task conditions in their study of local and global music perception that were directly comparable to the same/different judgments used by Heaton et al. and Mottron et al. They found no difference between their group of adolescents and adults with ASD and matched controls in detecting differences that either violated the pitch contour or left it unaltered. Their study did therefore not replicate the "local superiority" finding of Mottron et al.

Foxton et al. (2003) questioned, however, whether contour should be defined as the "global" level of a melody and proposed that rather than a high level of perceptual organisation, contour is simply a succession of perceived local features. They redefined pitch contour, such as change in pitch direction (up/down), as a "local" auditory feature, while the integration of contour with absolute pitch values and timing, form the "global" percept. Foxton et al. compared individuals with ASD with age- and IQ-matched control participants in their ability to detect whether melody pairs were equivalent in local pitch direction changes (i.e., whether the melodies follow the same pattern of rise and fall in pitch). The global structure of comparison sequences was altered by modifying several local features such as transposing the piece half an octave higher and varying the timing of the pitch direction change. Participants were therefore required to attend to a local feature while the extent of global interference was assessed (akin to the visual *global-to-local interference* effect with hierarchical stimuli). Changes at the global level were detrimental to task performance for control participants, but had less impact on the performance for participants with ASD. The authors attributed this finding to the absence of interference from the global level on auditory processing in autism.

2.5 Alternative explanations to weak central coherence

Empirical findings that are at odds with the original weak central coherence account, such as intact global processing (e.g., Heaton et al., 1999; Mottron et al., 2000; Plaisted et al., 1999), have led investigators to consider alternative explanations for the characteristic pattern of strengths and weaknesses seen in ASD. The three main alternative theoretical accounts addressing findings relevant to weak coherence are described below.

Plaisted (2001) has suggested that the perceptual and attentional abnormalities found in ASD stem from a reduced ability to generalise, that is, observe similarities between stimuli, which occurs alongside an enhanced ability to discriminate between highly similar stimuli. This theory developed from studies of perceptual learning (O'Riordan & Plaisted, 2001; Plaisted, O'Riordan, & Baron-Cohen, 1998a) and visual search (Plaisted, O'Riordan, & Baron-Cohen, 1998b), in which individuals with autism appeared to process unique features of stimuli well, but common features relatively poorly; opposite to what is observed in typical populations. This theory suggests that the mechanism underlying weak coherence operates at a perceptual level and, as a consequence, does not easily explain the high-level conceptual findings relevant to weak central coherence in ASD. In defence, Plaisted has argued that a reduced ability to process common features could also relate to difficulties in assimilating new information, which would impact on the ability to extract the gist or meaning from a situation.

Mottron and Burack (2001) have also suggested that low-level perceptual operations are atypical in autism in their 'enhanced perceptual functioning' account. They propose that the information processing systems employed in the detection, discrimination, and categorisation of perceptual stimuli are overdeveloped in individuals with ASD, with local bias seen as a side effect of this general enhancement. Excessive lower order perceptual functions are thought to interfere with the attainment of higher order processes, but are not necessarily associated with an imbalance between local and global processing. Deficits in global processing are therefore not predicted by this theory. Mottron and Burack claim the main strength of the model is that it provides "a neuro-cognitive-behavioural framework that can account for many of the characteristics of peaks of performance among persons with autism and that provides links between the developmental courses of cognitive operations and behaviours" (p. 145).

Baron-Cohen (2002) is the main proponent of the 'empathising–systemising' and 'extreme male brain' theories of autism. Based on an extensive literature on gender differences in cognition, Baron-Cohen proposed that individuals with autism show a profile resembling the extreme end of 'maleness'; that is, high performance on tasks where male superiority has been found (e.g., mathematical and mechanical reasoning, aspects of spatial visualisation such as mental rotation) and low performance on tasks where female advantage has been established (e.g., empathy, social judgment, decoding non-verbal communication, sensitivity to emotional expression).

Baron-Cohen (2002) characterised male and female brain types along dimensions of 'empathising' and 'systemising'. Empathising (including theory of mind) is defined as the drive to recognise emotions and thoughts in others and to respond to these spontaneously and appropriately. Systemising is defined as the drive to analyse and construct systems, and to extract underlying rules that govern the behaviour of a system. Individuals with ASD are suggested to show an extreme male profile with relatively high systemising and low empathising abilities. Empirical studies have so far shown support for this hypothesis. Self-ratings of abilities and preferences have confirmed that high-functioning adults with ASD have a stronger drive to systemise and a weaker drive to empathise than control participants (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003). Baron-Cohen, Wheelwright, Spong, Scahill, and Lawson (2001) found children with ASD to be superior to TD children on a test of folk physics (requiring understanding physical causality), while being impaired on a test of folk psychology (requiring understanding of social causality).

Lawson, Baron-Cohen, and Wheelwright (2004) also contrasted performance on tasks of systemising and empathising in high-functioning adults with either ASD or TD. Systemising was assessed through the ability to understand physical systems and predict movement in mechanical diagrams (the Physical Prediction Questionnaire; PPQ), while empathising was assessed by the ability to detect subtle *faux pas* presented in stories (the Social Stories Questionnaire; SSQ). Gender effects were found in TD adults, with females surpassing males on the SSQ and males performing higher on the PPQ. Males with ASD showed an extreme form of the male profile performing lower on the SSQ than the TD males, although not differing on the PPQ suggesting intact rather than superior systemising abilities. Empathising and systematising were considered to be independent constructs as no correlation was observed between the two measures.

Consistent with weak coherence theory, the 'empathising–systemising' model proposes that individuals with ASD are at the extreme end of a cognitive style that varies along a continuum in the general population. Baron-Cohen (2002) proposed that local processing was a prerequisite for systemising and explains the extreme detail-focus evidenced in ASD. A preoccupation with details may be necessary to begin the analysis of a system, however the goal of understanding the complexities of a whole system requires integration of several parts, which distinguishes this theory from weak central coherence. A strong point of the 'empathising–systemising' theory is, however, the capacity to explain why narrow but highly developed interests and preoccupations occur in individuals with ASD.

These alternative theoretical accounts of the findings underpinning the weak central coherence theory have stimulated much debate in the ASD research community. From this the assumption of an inherent trade-off between processing at the local and global levels has been revised (Happé & Frith, 2006), although the suggestion that weak coherence is a cognitive style or bias, such that global processing is possible, but local processing is the preferred option, is generally maintained.

2.6 The search for the mechanisms of central coherence

Although weak central coherence (and the alternative accounts) offers a viable explanation for some of the behavioural manifestations of ASD, the theory is still at a descriptive level and does not inform why or how such an effect occurs. Some speculations have been made to the possible underlying mechanisms of weak central coherence, at both cognitive and neural levels, and these are reviewed below. Of interest to the present study is whether proposed mechanisms suggest pervasiveness of coherence across levels and modality of cognitive functioning, or disparate performance in different domains.

Along with the alternative conceptualisations of central coherence findings, which cite the mechanisms at the perceptual level, some authors have proposed perceptual-attentional mechanisms are the root of detail-focus processing in ASD. Mann and Walker (2003) for example, found evidence of difficulty in broadening the spread of visual attention when making size judgments of cross-hair stimuli. A deficit was found when the successively presented stimuli changed from small to large, but not vice versa, suggesting a specific difficulty in "zooming out". However, an opposite effect was described by Burack (1994) who referred to the visual perception of individuals with ASD as an 'inefficient attentional lens' that was abnormally broad and not able to 'zoom in'. Burack found the facilitation effect of a window to direct attention to a target was lost in the presence of distracter stimuli, whether they were near or distant from the target.

In contrast to the majority of perceptual studies conducted in the visual modality, Plaisted, Saksida, Alcántara, Weisblatt (2003) investigated whether abnormalities in the peripheral processing of auditory stimuli could relate to perceptual differences in ASD. Using an auditory filtering task, they found that the width of auditory filters in autism were significantly wider than normal. This finding was contrary to their initial hypothesis that individuals with ASD might show greater than normal auditory frequency selectivity, reflecting an enhanced ability to focus on acute features at an early stage of perceptual processing. The broader auditory filters may, however, reflect the reported difficulty in ASD to filter out background noise.

2.6.1 *Computational models*

Several computational models have been developed that may aid in defining the cognitive and neural bases for weak coherence. In Cohen's (I. L. Cohen, 1994) neural network model, an excess of neuronal connections are said to exist in the brains of individuals with autism. In computational models this leads to a strong ability to recognise and discriminate stimulus patterns, but poor ability to generalise due to the over-emphasis on details; a pattern considered to be qualitatively similar to the learning and behavioural characteristics of autism. Along a similar argument, McClelland (2000) suggested the brains of individuals with autism are inclined to excessive conjunctive neural coding, a process that reduces the extent of overlapping patterns of activation in the neural circuitry. An unusual reliance on combinatorial coding is hypothesised to result in the development of semantic and conceptual representations that are highly specific and lacking in generalisation.

Gustafsson (1997) presented a neural circuit theory of autism that proposes impairment in the development of cortical feature maps, resulting in an inability to extract salient features from stimuli. In this model, neurons are arranged in columns so that adjacent columns become activated by objects with similar features. Gustafsson proposed that excessive inhibitory feedback in autism leads to an arrangement of feature maps in which these columns are too narrow, preventing the integration of information. Individuals with autism would be restricted to processing raw data, resulting in high sensory discrimination and abnormal sensory responses.

Finally, O'Loughlin and Thagard (2000) characterised coherence in terms of 'maximal constraint satisfaction.' In this model coherent elements are connected by positive constraints (excitatory links) and incoherent elements are connected by negative constraints (inhibitory links); coherence is operationalised as the drive to maximise the agreement of positive and negative constraints among elements. Weak coherence is viewed as a disruption to this balance, whereby the level of inhibition is greater than the level of excitation. This prevents the activation of coherent alternatives, often resulting in a local solution to a coherence problem. O'Loughlin and Thagard successfully applied this model to simulate weak coherence on the Homograph Reading task (Happé, 1997). It is not clear however, whether this model can account for local bias on visual-spatial tasks.

2.6.2 *The brain basis of coherence*

Several neurological studies have attempted to elucidate brain regions involved in local and global processing. Evidence that the right hemisphere is implicated in configural processing has emerged from studies of individuals with acquired right hemisphere damage (see Section 2.9). Neuroimaging studies using functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) methods also point towards hemispheric specialisation for global and local processing. Fink et al. (1997) for example, scanned TD adults while they attended to, or switched attention between, local and global levels of hierarchical figures. Areas of brain activation differed between the two levels: selective attention to global features was associated with right lingual gyrus activation, while attention to local features was associated with left inferior occipital activation. The act of switching between local and global levels resulted in further activation of the anterior cingulate and the dorsolateral prefrontal cortex and highlighted the role of executive attentional control in alternating attention between levels. The lateral dominance for local and global processing may not emerge until adolescence, as shown by fMRI studies of children analysing hierarchical patterns (Moses et al., 2002; see Chapter 3 for review of developmental aspects of local global processing).

It has been suggested that the right hemisphere may also mediate verbal coherence in high-level cognitive processes such as discourse comprehension. Robertson et al. (2000), for example, found that, compared to reading unrelated sentences, comprehending connected sentences produced greater activation in the middle and superior frontal regions of the right hemisphere. St. George, Kutas, Martinez, and Sereno (1999) also reported a greater degree of right hemisphere activation when participants read untitled stories relative to titled stories. It was suggested that the greater amount of effort required to create a coherent episode with an untitled story resulted in this greater activation. However, contrary findings regarding the role of the right hemisphere in verbal coherence have been reported. Ferstl and von Cramon (2001), for example, asked participants to judge whether two successively presented sentences were pragmatically coherent or not, and found no evidence for right hemisphere involvement in making inferences (see Gernsbacher & Kaschak, 2003, for a review of neuroimaging work on language processing).

To date, there is little evidence for localised abnormalities in the brains of individuals with ASD that relate to local-global processing. A few studies have reported right hemisphere dysfunction in case studies of Asperger syndrome (McKelvey, Lambert, Mottron, & Shevell, 1995) and reduced white matter volume in the right hemisphere in adolescents and young adults with ASD (Waiter et al., 2005). Asperger syndrome has also been likened to nonverbal learning disabilities syndrome (Rourke, 1995), which has been suggested to result from dysfunction of white matter affecting right hemisphere functioning and interhemispheric communication (Gunter, Ghaziuddin, & Ellis, 2002). However, as reviewed by Bauman and Kemper (2005) and Brambilla et al. (2003), many other brain regions have been implicated in ASD that are non-specific to the right hemisphere (e.g., the limbic, frontal and cerebellar regions), and it is as yet unclear which anomalies are specific or universal to ASD.

There has been one functional imaging study of coherence in individuals with ASD. Ring et al. (1999) questioned whether superior ability on the Embedded Figures Test (EFT) in ASD was due to distinct differences in brain activation during task performance. They found that searching for hidden figures activated similar cerebral regions in adults with ASD and TD adults, which had previously been implicated in object and visual-spatial processing, although some differences suggested that the cognitive strategies employed distinguished the two groups. Greater frontal and parietal activity was observed in the TD adults, while greater activation of occipital regions was found in the ASD group. This suggested that early stages of sensory processing, such as object feature analysis, were being utilised in ASD. In contrast, contributions from high-level visual perception, such as top-down modulation (necessary to extract global features) were not functioning to the same degree in ASD as compared to TD adults during embedded figures detection. The difference in high- and low-level processes in embedded figure detection may concur with the recent findings of Jarrold, Gilchrist, and Bender (2005). Embedded figures detection in children with autism was related to simple visual search for targets identified by a unique feature. In contrast, embedded figure detection was specifically associated with visual search for targets identified by a conjunction of features in TD children. This dissociation may reflect different processes used in embedded figure detection where low-level processes are employed in ASD and more high-level processes are implicated in typical development.

The absence of a well-matched control task limited the conclusions that could be made by Ring et al. (1999). In response, Manjaly et al. (2003) isolated the disembedding process by contrasting performance on embedded versus non-embedded versions of the task. It was revealed that regions in the left inferior and superior parietal cortex and left ventral premotor cortex were specific to searching for embedded figures in TD individuals. As these regions are close to posterior language areas of the left hemisphere, the findings were discussed in relation to the strong association between aphasia and impairment on the EFT (Teuber & Weinstein, 1956). Future studies with volunteers with ASD are planned.

2.6.3 Neural pathways involved in coherence

As well as the study of particular regions of the brain, abnormalities are also suggested in specific neural pathways that may underlie weak coherence in ASD. Recent interest in the perceptual and motor related abnormalities reported in ASD has led to investigations of possible impairments in low-level visual processing, specifically in the magnocellular (dorsal stream) visual pathway. This process is commonly assessed by ability to detect coherent motion from an array of randomly moving dots. Individuals with ASD have demonstrated high motion coherence thresholds, implying an impaired ability to integrate moving local visual signals to extract a global pattern (Milne et al., 2002; Spencer et al., 2000). Milne (2003) also found that a subgroup of children with ASD who had a specific difficulty in detecting coherent motion also demonstrated weak coherence as shown by superior performance in the EFT and a local bias when processing Navon hierarchical figures. Pellicano, Gibson, Maybery, Durkin, and Badcock (2005) replicated this finding and reported a significant inverse correlation between global motion perception and detecting embedded figures in a group of children with ASD, suggesting that the mechanism of weak visuo-spatial coherence may be directly linked to the magnocellular pathway. Although global motion detection was impaired, the children with ASD in the Pellicano et al. study demonstrated normal flicker contrast sensitivity, suggesting lower-level dorsal stream functioning was intact. Bertone, Mottron, Jelenic, and Faubert (2003) also reported normal first-order (luminance-defined) motion perception, but not second-order (texture-defined) motion perception, in individuals with autism. The authors suggest that visual processing impairments in ASD may therefore be at a complex neural level, when integrative processing is required.

2.6.4 *Neuronal connectivity models of coherence*

Computational models have suggested that the processes underlying neural integration are impaired in autism, which leads to deficits associated with weak central coherence (see Section 2.6.1). There is also considerable concrete evidence for abnormal brain connectivity in ASD from neurological studies (see Belmonte et al., 2004, for a review). A plausible explanation is that specialised brain regions that typically connect and collaborate while performing a task, may be less connected and work in isolation in the brains of individuals with ASD. This failure to integrate information between brain regions is suggested to be indicative of weak central coherence.

Brock, Brown, Boucher, and Rippon (2002) have presented one such theory of under-connectivity in their 'temporal binding deficit' hypothesis of ASD. Temporal binding refers to the synchrony required by specialised neurons as they respond to different elements of the same scene and, as a consequence, bind parts into wholes. This process has been demonstrated by studies showing temporally correlated EEG activity at high frequencies (gamma-band) during normal visual perception of coherent objects. If individuals with ASD perceive objects as a collection of parts rather than as a whole, a reduction in this synchronisation is predicted. Brock et al. speculate similar effects would occur in verbal-semantic coherence, where a deficit in this binding process would result in the processing of individual words independent of context. This hypothesis has yet to be tested in ASD although the authors suggest potential methods using EEG techniques.

Just, Cherkassky, Keller and Minshew (2004) also propose that reduced synchronisation between cortical areas, as found in their fMRI study, might account for the lack of integrative processing in ASD. Individuals with and without ASD were scanned as they responded to comprehension questions about simple sentences they read. Although both groups activated the classic language areas while performing the task, the ASD group produced more activity than the control group in Wernicke's area (left latero-superior temporal), and less activity in Broca's area (left inferior frontal gyrus). This pattern of activation in the ASD group suggested a greater degree of low-level processing (i.e., of individual words) and less high-level integration of meaning and working memory as required in sentence processing. Just et al. also reported that the degree of connectivity between cortical language regions, as assessed by the amount of correlation in activation, was reduced in ASD compared to control participants. Reduced connectivity between brain regions in ASD as participants watched non-verbal animations requiring attribution of intention was previously reported by Castelli, Frith, Happé, and Frith (2002). Just et al. argue that their underconnectivity theory provides a platform for the underlying biological structures and processes of weak coherence in autism.

Several researchers have speculated how reduced connectivity may occur in the brains of individuals with autism. The most simplistic account would be a lack of connecting fibres, possibly resulting from excessive pruning of connections. Chung, Dalton, Alexander, and Davidson (2004) for example, found reduced white matter density in the corpus callosum in individuals with HFA suggesting impaired interhemispheric connectivity (also confirmed by Waiter et al., 2005). Barnea-Goraly et al. (2004) also found reduced white matter integrity in adolescents with ASD compared to controls using diffusion tensor imaging (DTI).

These findings may be associated with unusual growth patterns during critical periods (Courchesne et al., 2001). There is substantial evidence for increased brain size, particularly in cerebral white matter, during early childhood in ASD compared to TD children (see Chapter 1, Section 1.4.2). Frith (C. Frith, 2004) proposes that a lack of pruning during the normal growth spurt may result in the presence of unnecessary connections and cause increased brain size. This process would likely impact on the connectivity between brain regions. A lack of neuronal pruning of feedback connections would also affect top-down processing systems, which directly relate to integration and coherence processes (Hill & Frith, 2003). A disruption in the connectivity between brain regions is not unique to ASD and has been suggested in relation to other clinical disorders, such as schizophrenia and dyslexia (Friston, 1999; Paulesu et al., 1996), and thus limits the specificity of this account. The theory does have intuitive appeal however, as it proposes a viable link between biological and cognitive models.

2.7 The universality of weak coherence in ASD

A challenge to the validity of weak central coherence theory is the extent to which this processing style can be said to characterise all individuals with ASD. It has been argued that weak central coherence should be present universally if regarded as a causal mechanism in the development of the disorder (Pellicano, 2004). However, similar to the finding that impairments in theory of mind and executive functioning do not characterise all individuals with ASD, it is accepted that weak coherence may only be present in a proportion of those with the disorder. Furthermore, developmental change or compensation may disguise cognitive style in some individuals with ASD.

Several studies have reported the degree of universality of weak coherence in their studied group of individuals with ASD. Happé (1994b) reported that 85% (41 from 48) of her sample of children and adults with autism were found to have peak performance on the Block Design subtest of the Wechsler scales, relative to their own mean performance scaled score. Pellicano (2004) also found high proportions of children with ASD showing weak coherence, defined as performance greater than one standard deviation from the mean of the TD group. Using this criterion, 92% (48 from 52) of children with ASD demonstrated quick performance on the EFT, 73% (38 from 52) obtained high scores on the Pattern Construction test, and 77% (40 from 52) performed well on the Figure-Ground task. When scores were adjusted to account for lower verbal abilities of the ASD group, all percentages exceeded 85% demonstrating the high prevalence of local processing abilities in her sample.

In contrast, Teunisse, Cools, van Spaendonck, Aerts, and Berger (2001) found that weak coherence was not a universal feature in their sample of high-functioning adolescents with autism. They report that 57% (20 from 35) of their group of high-functioning adolescents with autism performed below one standard deviation from the mean of the normative data on the Silhouette subtest from the Visual Object and Space Perception Test. A similar proportion was reported in Jarrold and Russell's (1997) study of counting styles, where 55% (12 from 22) of children with autism were found to show a local cognitive style as demonstrated by a lack of benefit from canonical patterns. This was in contrast to 5% (1 from 22) TD children, but not significantly different from 32% (7 from 22) of children with MLD who also showed this pattern of performance. Loth, Gómez, and Happé (2006) also examined the universality of weak coherence in ASD. According to the pattern of individual performance across three coherence measures (EFT, Block

Design, Sentence Completion), 47% (22 from 47) of individuals with ASD showed weak central coherence, compared to 6% (2 from 32) of TD children and adults. No individuals with learning difficulties (N = 14) were observed to show this processing style.

Age, intelligence, gender and diagnosis may all be factors relevant to individual differences in weak coherence in the ASD population. Studies that have included participants at the high ability end of the spectrum, for example, have not reliability found the predicted result (e.g., Burnette et al., 2005; Ropar & Mitchell, 2001), which suggests that this processing style may be more prevalent in low-functioning individuals. The existence of individual differences in coherence reflects the vast heterogeneity of individuals placed on the autism spectrum. It is hoped that these individual differences may in the future allow identification of meaningful subgroups within the ASD population, which may differ in manifestation, prognosis, or response to remediation.

2.8 The explanatory power of weak central coherence

When Frith (1989) proposed the theory of weak central coherence she suggested that it may describe the primary cognitive deficit in autism and could even explain the prominent social impairments of the disorder. Since this proposal there has been much debate over how much weak central coherence can account for the symptomatology of autism and whether this theory is independent from other established theories of autism such as theory of mind and executive dysfunction. The relationship between these processes has been scrutinised in individuals with ASD as well as those with typical development and, as shown in the following review, the association is not conclusive.

2.8.1 Weak central coherence and theory of mind

Several aspects of social communication require the ability to integrate different parts of a social situation. Frith (1989), for example, observed "ordinary conversation and the understanding and answering of questions as intended by the questioner implies striving for high-level global, not merely local, coherence of information" (p. 101). Detail-focused processing at a perceptual level may also impact on social learning. Children with autism are thought to process faces in terms of individual features, rather than their overall configuration (Hobson, Ouston, & Lee, 1988). As many emotions are distinguishable only when different facial features are integrated into a whole, this important area of social perception is likely to be affected by weak coherence.

The plausible link between social impairments (as indexed by a lack of theory of mind) and weak central coherence was not supported by empirical studies that demonstrated their independence. Weak coherence, as shown by superior Block Design (Happé, 1994b) or insensitivity to context when reading homographs (Happé, 1997), was present in individuals with autism irrespective of their performance on theory of mind tasks. In a revision of the weak central coherence account, Frith and Happé (1994) suggested that two independent processes might coexist in autism, namely, a deficit in mind-reading ability and weak central coherence.

Challenges are still being made to this position however. Jarrold, Butler, Cottington, and Jimenez (2000), for example, reported an inverse relation in a sample of undergraduate students between performance on the Eyes task (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997), in which mental states have to be interpreted from photographs of the eye region, and speed on the EFT. The finding that poorer theory of mind performance was associated with weaker central coherence was also confirmed in TD children and children with autism. Jarrold et al. reported that theory of mind ability (as assessed by false belief test performance) was inversely related to performance on two measures of central coherence, the Pattern Construction subtest of the Differential Ability Scales (Elliot, 1990) and the preschool version of the EFT (Coates, 1972), in a group of 5-year-old TD children, once verbal mental age (VMA) was taken into account. Furthermore, this association was replicated in a group of children with autism (aged 7 to 12 years), but again, only after controlling for differences in VMA.

Pellicano, Maybery, and Durkin (2005) were not able to replicate the findings of Jarrold et al. (2000). Superior theory of mind performance was associated with faster detections on the preschool version of the EFT and higher Pattern Construction scores in their large sample of 4- and 5-year-old children. The effect of partialling out individual effects of age, verbal and non-verbal ability rendered the correlations non-significant, and provided no evidence that weak coherence was associated with poor theory of mind. This lack of association between the two processes was also found in children with ASD (Pellicano, 2004). Although poor performance on the theory of mind tasks was significantly related to local processing on the visuo-spatial measures of coherence, this association dissipated once the effects of age, verbal ability and non-verbal ability were removed.

Pellicano's (2004) findings are also consistent with Morgan, Maybery and Durkin (2003) who investigated the explanatory power of weak central coherence to suggested precursors of theory of mind: joint attention and pretend play. Twenty-one preschool children with ASD (aged 3 to 5 years) were compared to 21 control children matched in age and nonverbal ability. As predicted by central coherence theory, the children with ASD were significantly quicker than control children in identifying embedded figures and constructing block patterns. Furthermore, the children with ASD made significantly less eye contact during interactions with the experimenter and, despite showing a similar amount of pretend play, demonstrated less functional play than control children. These differences were not explained by the lower VMA of the children with ASD and no association was found between the measures of central coherence, joint attention, and VMA. In fact, all three domains made significant and independent contributions to predicting group membership, prompting the authors to suggest that three independent cognitive deficits may underlie autism.

Burnette et al. (2005) recently reported that insensitivity to sentence context, as indicated by the number of homographs incorrectly pronounced, was related to poor performance on first-order and second-order theory of mind tasks in a group of high-functioning children with autism. This relationship remained marginally significant after verbal intelligence was controlled for, suggesting that verbal skills were not contributing to this association. No relationship was found between the visuo-spatial measures of coherence and theory of mind skills, suggesting that the processing domain selected to assess central coherence may be of relevance. Taken together, the reviewed studies provide mixed evidence regarding whether central coherence and theory of mind are dissociable processes.

2.8.2 *Weak central coherence and executive dysfunction*

As outlined in Chapter 1, the executive dysfunction account of autism has been successful in addressing the non-social deficits of ASD, most notably the characteristic rigid and repetitive behaviours. It has been questioned whether findings currently attributed to weak central coherence could, in fact, be explained by executive dysfunction. It is easy to see how weak central coherence could be interpreted as a consequence of impaired executive control. Hoy et al. (2004) for example, proposed that failure to select the correct interpretation of a homophone that has been disambiguated by context could arise from a lack of shifting from the original interpretation of the word. Indeed, one theory of poor reading comprehension suggests an inability to suppress irrelevant information (Gernsbacher, Varner, & Faust, 1990). Poor inhibition might explain why in ASD the frequent pronunciation of a homograph is given rather than the context-appropriate rare pronunciation (Happé, 1997). Other facets of executive dysfunction, such as limited working memory, poor planning, or inability to shift between local and global dimensions of stimuli, might also result in patterns of performance that are interpretable as local bias (Happé & Frith, 2006).

Three recent studies have investigated the possible link between central coherence and executive functions. Booth et al. (2003) provided evidence that weak coherence was not accounted for by one aspect of executive dysfunction: poor planning skills. A novel approach was used whereby both central coherence and planning ability were assessed within the same drawing task. In order to examine the effect of poor executive functioning on coherence, boys with Attention Deficit-Hyperactivity Disorder (ADHD) were included. Participants were asked to copy simple objects (e.g., a snowman) and then redraw the object with an added internal feature (e.g., a snowman with teeth) that necessitated planning ahead for its inclusion (e.g., larger head size). While both ADHD and ASD groups showed some planning deficits in their attempts to include the extra feature, only the boys with ASD showed a fragmented, piecemeal drawing style. The two processes of planning and drawing style were unrelated in the clinical groups, although were associated in a TD sample, such that poor planners were *less* likely to be detail-focused in their drawing style. It was concluded that poor planning was not able to explain detail-focused processing in ASD.

Pellicano, Maybery et al. (2005) reported an association between performance on visuo-spatial measures of coherence and executive functioning in their sample of 4- and 5-year-old TD children. Good performance on both the Pattern Construction and the Developmental Test of Visual-Motor Integration was reliably related to proficient planning, flexible thinking, and working memory capacity. However, this suggested that visuo-spatial ability, rather than strong central coherence, was associated with good executive control. Pellicano (2004) also confirmed that superior visuomotor integration was associated with better inhibition, planning and cognitive flexibility in high-functioning children with ASD and TD controls. Tasks requiring local processing skills did not show any relation to executive control.

Teunisse et al. (2001) concluded that no relationship existed between weak central coherence and poor cognitive shifting in a sample of high-functioning adolescents with autism and a comparison group of TD individuals. Furthermore, both processes were found to be unrelated to measures of symptom severity or social competence. However, in a three-year follow-up of these participants, Berger, Aerts, van Spaendonck, Cools, and Teunisse (2003) reported that degree of cognitive shifting, rather than central coherence, predicted good outcome. Individuals with autism who performed relatively well on tasks such as the Wisconsin Card Sorting Test (WCST) and the Intra-dimensional/Extra-dimensional shift task (ID/ED, from the Cambridge Neuropsychological Test Automated Battery; CANTAB), showed greater improvement in social competence as assessed by the Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cichetti, 1984) over the three-year interval. Longitudinal studies may therefore prove to be extremely helpful in investigating the possible causal relations between cognitive processes underlying autism.

2.9 Specificity of weak central coherence

A weaker drive for central coherence appears to be specific to individuals with ASD, although comparisons are often only made to TD participants. Alternative control groups have been used in a few studies, such as individuals with mental retardation (Jarrold & Russell, 1997; Tager-Flusberg, 1991), learning disorders (Hermelin & O'Connor, 1970), language disorders (Norbury & Bishop, 2002), Tourette syndrome (Ozonoff et al., 1994), and ADHD (Booth et al., 2003; Happé & Booth, 2006), with the specificity of weak coherence to ASD generally confirmed. Local-detailed processing has, however, been empirically shown in other clinical populations, such as schizophrenia, mood disorder, right hemisphere damage, and Williams Syndrome; and a brief review of these findings follows.

Abnormal attention and visual perception have long been considered core deficits of schizophrenia and research suggests that this disorder may be associated with more efficient processing of local details, relative to the gestalt. Ferman, Primeau, Delis, and Jampala (1999), for example, reported that patients with schizophrenia showed a local advantage as well as a global disadvantage when responding to hierarchical figures, contrary to typical populations. Chen, Nakayama, Levy, Matthysse, and Holzman (2003) also reported that individuals with schizophrenia showed a selective deficit in detecting coherent motion of random dots, which requires global processing of the target. Furthermore, Uhlhaas, Silverstein, Phillips, and Lovell (2004) found that non-clinical groups high in schizotypy showed reduced use of visual context and did not readily succumb to visual illusions.

Mood states are known to affect processing biases, with negative states relating to more detail-focused, analytical processing and less global processing (Gasper, 2004). This has been found in individuals high on trait anxiety and depression (Basso, Schefft, Ris, & Dember, 1996; Derryberry & Reed, 1998). Even inducing a negative mood in a testing situation can influence processing style towards local details over the global form (Gasper & Clore, 2002). These findings are of relevance to the weak central coherence theory as elevated levels of anxiety and depression are commonly found in ASD, notably at the high ability end of the spectrum. This has been confirmed in children with Asperger syndrome as assessed by parental report (Kim, Szatmari, Bryson, Streiner, & Wilson, 2000) and in children with HFA using a self-report measure (Gillott, Furniss, & Walter, 2001). It remains to be seen whether negative mood is related to local-detailed processing in individuals with ASD. One study to date has tested this hypothesis (Burnette et al., 2005), but found no relationship between self-report measures of anxiety and depression and performance on a range of coherence tasks in children with HFA. It is not yet known, however, whether self-report measures of internal states are reliable and valid indicators in individuals with ASD.

Neuropsychological studies of patients with lateralised brain damage show differential patterns of local and global processing. Specifically, patients with right hemisphere damage have displayed impairments in identifying global level forms, while patients with left hemisphere damage have shown deficits in the identification of the local level elements (Delis, Kiefner, & Fridlund, 1988; Lamb, Robertson, & Knight, 1990; L. C. Robertson, Lamb, & Knight, 1988). Children with right hemisphere congenital brain injury also show deficits processing configural form (Stiles-Davis, Janowsky, Engel, & Nass, 1988). In free drawing tasks, these children were observed to draw individual features, but the picture lacked overall organisation and cohesion. In addition to visuo-spatial deficits, damage to the right hemisphere has also been found to impair the ability to integrate verbal information (Benowitz, Moya, & Levine, 1990). Benowitz et al. also found cross-domain deficits in coherence in right hemisphere patients and suggested "appreciation of spatial configurations and the comprehension of interrelationships among elements in narrative material may to some extent require a common mechanism" (p. 240).

The spatial-cognitive deficits observed in right hemisphere damage are reminiscent of the unusual feature processing styles seen in Williams Syndrome (Bellugi, Lichtenberger, Jones, Lai, & St. George, 2000). Individuals with this rare genetically-based disorder have a marked local bias alongside an almost complete disregard for the global form (Bihrle, Bellugi, Delis, & Marks, 1989). Surprisingly, the area of face processing is relatively spared in Williams Syndrome suggesting dissociation between spatial cognition and the speciality of face processing (Bellugi et al., 2000).

It is questionable whether the underlying cognitive style presented in these groups is same as in ASD. It is clear that further study is required where coherence task performance in these different clinical groups is directly compared with individuals with ASD before any conclusions can be made.

2.10 Weak central coherence in the broader autism phenotype

Evidence of cognitive biases, akin to weak central coherence, has been found in relatives of individuals with ASD. Baron-Cohen and Hammer (1997) found that parents of children with autism showed superior performance on the EFT, while also being impaired on a task related to theory of mind (the Eyes task). Happé, Briskman and Frith (2001) performed a comprehensive assessment of central coherence in parents of boys with autism. They found that approximately half of the fathers and a third of the mothers of boys with autism showed consistent weak central coherence across the test battery, which was not evident in control groups of parents of boys with dyslexia and boys with typical development. Interestingly, cross-domain coherence was found where weak coherence in the visuo-spatial domain (i.e., fast and accurate performance on the Block Design and EFT) was coupled with weak coherence in the verbal-semantic domain (i.e., completing sentence stems with a local associate that was meaningless to the context).

Support for weak central coherence as an indictor for the broader phenotype of autism also emerges from everyday processing preferences and choice of occupation in family members. Baron-Cohen, Wheelwright, Stott, Bolton, and Goodyer (1997) found a higher than expected proportion of fathers of children with autism were engineers, a profession where attention to detail can be advantageous. A further study confirmed that autism was more common in the families of engineers, physicists, and mathematicians (Baron-Cohen et al., 1998). Briskman, Happé, and Frith (2001) found that self-report measures of everyday weak coherence, such as detail-focused interests and abilities (e.g., noticing small errors), related to weak central coherence on experimental measures in parents of boys with autism.

2.11 The pervasiveness of weak central coherence

As reviewed in Section 2.4, ASD has been characterised by weak central coherence at a number of levels: perceptual, auditory, visuo-spatial, and verbal-semantic. It remains to be seen whether this processing style shows consistent individual differences *across* levels and domains. Fundamentally, performance on all coherence tasks should correlate if the same construct is being measured. The following provides a review of studies that have administered a battery of tasks to obtain a more descriptive analysis of weak central coherence and to determine whether the same underlying construct is assessed in various tasks purporting to measure local-global processing.

2.11.1 Within-domain findings

Five studies are presented that specifically examine the validity of the central coherence construct by administering a battery of coherence measures within the same processing domain. All but one study (Jolliffe & Baron-Cohen, 1999) investigated the convergent validity of coherence in the visuo-spatial modality.

Ropar and Mitchell (2001) compared performance across three visuo-spatial tasks that have previously been implicated in studies of weak coherence in autism: the EFT, the Wechsler Block Design subtest, and the Rey Complex Figure (accuracy in copy and immediate recall). The task battery was administered to children with autism (N = 19, aged 9 to 18 years), Asperger syndrome (N = 11, aged 8 to 15 years), MLD (N = 20, aged 9 to 14 years), and two groups of TD children, aged 8 years (N = 19) and 11 years (N = 18). Strong correlations were found in all groups between the visuo-spatial measures, which generally remained after the effects of age and VMA were removed (with the exception of the MLD group and the youngest TD group). General visuo-spatial abilities were not controlled for however, and it is possible that these skills were driving the association between measures rather than central coherence. For example, it is questionable whether accuracy in copying the Rey Complex figure is a valid assessment of coherence as a piecemeal or fragmented approach can result in a globally coherent drawing (Booth et al., 2003), although this drawing style may impede accurate recall (Akshoomoff & Stiles, 1995b).

Ropar and Mitchell (2001) also investigated the link between performance on the battery of visuo-spatial tasks and degree of susceptibility to different visual illusions. Only in 11-year-old TD participants, did good performance on visuo-spatial tasks relate to less susceptibility to one illusion (the Muller-Lyer illusion). Conversely, in individuals with autism and Asperger syndrome, good performance on the visuo-spatial tasks predicted a greater degree of susceptibility to two visual illusions (the Hat and Titchener circles); a finding that was not in line with predictions of weak central coherence theory. Ropar and Mitchell concluded that, overall, performance on the visuo-spatial tasks did not predict susceptibility to illusions, and suggested that different mechanisms may be in operation. However, the role of general or spatial intelligence cannot be distinguished from the role, if any, of cognitive style in this study.

Mottron, Burack, Iarocci, Belleville, and Enns (2003) attempted to isolate the particular aspect of local-global processing that was deficient in autism. This was achieved by administering a battery of visual coherence tasks that tapped different processing demands: local processing, global processing or switching between the two levels. Participants included 12 individuals with HFA (aged 9 to 22 years) and 12 age- and ability-matched control participants. No group differences were found on three measures of global processing: identifying fragmented versus complete letters, identifying silhouetted versus detailed objects, and visual search for targets identified by global grouping. Furthermore, groups did not differ on a divided attention version of the Hierarchical Figures task, in which target letters could appear at either the local or global level (although the typical global advantage was not observed in the control group on this task). Only on a measure of local processing did group differences occur. TD participants took longer to identify embedded versus isolated stimuli, but search times were similar for participants with HFA between the two conditions. Mottron et al. suggest that this provides evidence that autism is associated with local processing bias that is not necessarily a consequence of, or accompanied by, a global processing deficit. Group effects only were reported in this study, however, and relationships between measures were not given which may have provided more information on the dimensionality of coherence.

Pellicano, Maybery et al. (2005) questioned whether central coherence was a unitary construct in a group of 4- and 5-year-old TD children (N = 76). Four measures intended to tap coherence in the visuo-spatial domain were administered: the pre-school version of the EFT (Coates, 1972), the Figure-Ground task from the Developmental Test of Visual Perception (Hammill, Pearson, & Voress, 1993; assessing the ability to detect several shapes embedded in a complex background), the Pattern Construction task from the Differential Ability Scales (Elliot, 1990; akin to the Wechsler Block Design subtest), and the Developmental Test of Visual-Motor Integration (Beery, 1997; assessing the ability to maintain form and spatial relationships when copying geometric figures). Contrary to expectations, correlations between all four measures were not as predicted by weak central coherence theory and a single coherence factor was not found. An exploratory factor analysis instead revealed a two-factor solution, the first of which incorporated the Pattern Construction and the Developmental Test of Visual-Motor Integration tasks. Good visual integration in copying geometric figures was related to quick and accurate performance on the Pattern Construction task suggesting this reflected visuo-spatial construction ability. This is in contrast to weak central coherence theory, which would have predicted a compromise between local (on Pattern Construction) and global (on the Developmental Test of Visual-Motor Integration) processing as evidenced by a negative correlation. The pre-school EFT and the Figure-Ground loaded on the second factor, but were in opposing directions: that is, fast identifications on the EFT were associated with fewer hidden shapes detected on the Figure-Ground task. This surprising result was thought to reflect search strategy rather than proficiency in analytic processing skills. This pattern of results was found even after differences in age, verbal and non-verbal ability were controlled for. Pellicano et al. concluded that central coherence in TD preschoolers is not a unitary construct.

Pellicano (2004) administered the same battery of visuo-spatial coherence tasks as described in Pellicano, Maybery et al. (2005) to 52 children with ASD (aged 4 to 7 years) and to a comparison group matched in age and nonverbal ability, although with higher verbal ability than the ASD group. The children with ASD demonstrated evidence of local processing across all coherence measures, as shown by faster detection times on the EFT, higher scores on the Pattern Construction and Figure-Ground tasks, and lower scores on the Developmental Test of Visual-Motor Integration. These differences were still evident after differences in verbal ability had been adjusted for. Raw correlations between the coherence measures were significant within each group and generally confirmed predictions from weak coherence theory. The exception was the positive association between the Pattern Construction task and the Developmental Test of Visual-Motor Integration (as found in the TD study, Pellicano, Maybery et al., 2005), which appeared to reflect the shared demands of visuo-spatial integration abilities. Adjusting for the effects of age and ability removed the significance of this correlation in the ASD group, although it remained in the control group. Most relationships between tasks disappeared after the effects of age and ability were partialled out from analyses, with exception between EFT and Pattern Construction in the ASD group. Pellicano suggested that, although at a group level weak coherence could differentiate between ASD and control children, at an individual level associations between coherence measures could not be explained beyond the effects of age, verbal ability, and nonverbal ability. However, weak coherence may be a defining feature of autism and show little within-group variability. Correlations may be low in a group selected for the key trait as a consequence of limited variance.

In contrast to the reviewed studies documenting the convergence of coherence on visuo-spatial tasks, Jolliffe and Baron-Cohen (1999) explored coherence in the verbal domain. High-functioning adults with autism or Asperger syndrome were impaired in their use of sentence context to determine the correct pronunciation of homographs or to interpret an ambiguous sentence read out to them. They also had difficulty with inferential reasoning when required to select the most appropriate bridging sentence. Of the two clinical groups, the autism group were found to have greater difficulty in achieving coherence. The three verbal coherence tasks were found to correlate with each other (and generally upheld partialling out IQ), suggesting that a unitary force may have driven performance, possibly weak central coherence.

2.11.2 Between-domain findings

Five studies to date have investigated the presence of weak central coherence across different processing domains. Although the first two studies do not compare individual performance across tasks, they have been included in this review since coherence measures from both visual and verbal domains were administered.

Lopez and Leekam (2003) assessed the degree of contextual processing in visual and verbal domains in a group of 15 high-functioning children with autism and 16 age- and IQ-matched TD children (mean age 14 years). Both groups were found to benefit from meaning when recalling semantically related and unrelated items that were presented verbally and pictorially. Priming effects from a word or a scene were also present in both groups. Group differences were detected only on a sentence-processing task where the children with autism were less likely to use sentence context to disambiguate homographs. Although individual performance across domains was not examined, Lopez and Leekam conclude that children with autism are facilitated by context information presented either visually or verbally and can make meaningful connections between single items. They suggest that the difficulty in using sentence context to disambiguate homographs may reflect a specific impairment in integrating multiple items of information.

Hoy et al. (2004) investigated whether weak central coherence was evident in both visual and verbal domains in a group of 17 children with autism (aged 6 to 9 years) and a group of TD children, matched in age and verbal ability to the autism group. A homophone task was employed as the verbal measure. Participants listened to sentences that contained a homophone (e.g., "The sale/sail was very big"). A second sentence followed that disambiguated the homophone either towards its frequent ("Everything was half price") or rare ("It made the ship go very fast") interpretation. Participants were then asked to select a picture that best showed the meaning of the homophone. Children with autism made significantly more errors when selecting the correct picture than control children, most notably when the rare interpretation was implied by the context. The group difference was, however, accounted for by individual variations in verbal ability. The set of visual illusions and associated control stimuli as used by Happé (1996) were taken as the visual measure of central coherence. No differences were found between children with autism and TD children on this task, with both groups succumbing to an equal number of illusions. Hoy et al. conclude that their study provided little support for weak central coherence in autism, or as a cross-domain phenomenon. Direct comparisons between the two tasks were not provided however, which might have elucidated any individual differences in coherence across visual and verbal processing domains. The authors suggest that the selected tasks may not have been adequate measures of central coherence, suggesting, for example, that the set of homophones were too difficult for participants. It was not stated, however, whether errors made to rare homophones were general errors, or responses made to the more frequent interpretation of the word, which would suggest a disregard for context. It is clear that a range of measures tapping central coherence in verbal, auditory and visual domains is needed to conclude whether this processing style has a cross-domain influence.

Teunisse et al. (2001) administered a comprehensive battery of coherence measures in their study of 35 high-functioning adolescents with autism (aged 16 to 24 years, no control group). An exploratory factor analysis determined that weak central coherence was captured by two distinct factors: piecemeal processing on visual-perceptual tasks and poor processing of meaning, in both visual-perceptual and verbal domains. The first factor included both child and adult versions of the EFT and a novel "search-for-difference" task where participants were required to compare two meaningful drawings and detect small differences in detail. The second factor comprised three object-recognition tasks from the Visual Object and Space Perception Test (Warrington & James, 1991), all testing the ability to identify objects based on obscured or degraded information (such as a rotated silhouette) and thus requiring strong coherence to deduce meaning from the limited visual information given. The tendency to recall words in the presented order and not use semantic clustering strategies on the California Verbal Learning Task (Delis, Kramer, Kaplan, & Ober, 1987) was also incorporated in the second factor. Therefore, a difficulty in perceiving meaningful form on the three visual-perceptual tasks was associated with less use of meaning to organise verbal recall. Teunisse et al. suggest that both factors of central coherence, piecemeal processing and ability to process meaning, need to be assessed to conclude that an individual exhibits weak central coherence.

Burnette et al. (2005) recently assessed the viability of the central coherence construct by administering a battery of coherence tasks to a group of 23 children with HFA and 20 age- and ability-matched control children (aged 8 to 13 years). Performance on three established visuo-spatial measures of coherence (EFT, Pattern Construction subtest from the Differential Abilities Scale, Wechsler Block Design subtest) were highly correlated in each group. Central coherence was not demonstrated across domains however as sensitivity to context on the Homograph Reading task (Happé, 1997) correlated with good performance on the visuo-spatial measures. Thus, contrary to predictions, strong verbal coherence was associated with weak visual coherence. The authors did not control for the individual effects of IQ however, which may underlie the association and mask real effects of cognitive style. Furthermore, children with HFA did not show the expected group difference on the visuo-spatial measures of coherence; surprisingly, the control children showed significantly higher performance on the EFT and the Pattern Construction subtest. Group differences were detected however on the Homograph Reading task with the children with autism less likely to provide appropriate pronunciations of homographs presented in a sentence context than control children. With the acknowledgement of individual differences in this processing style, this study suggests that weak central coherence in the visuo-spatial domain may not be characteristic of this sample of children with autism.

Loth et al. (2006) also questioned the prevalence of local-global processing style across visuo-spatial and verbal-semantic domains. Participants included 47 children and adolescents with ASD (aged 8 to 28 years) and an ability-matched control group. Central coherence in the visuo-spatial domain was indexed by Block Design and EFT performance, while the Sentence Completion task was used to indicate weak central coherence in the verbal-semantic domain. Cross-domain coherence was evident in the participants with

ASD: quicker times on the EFT were associated with more local completions on the Sentence Completion task. This relationship held even after the effects of age and IQ were controlled for. This correlation was not evident in the TD participants and was in the opposite direction in the group with MLD, where poor performance on the Block Design subtest (global bias) was related to more local completions made on the Sentence Completion task (local bias). The distinct pattern of detail-focused processing in visual-spatial tasks corresponding with degree of local processing in the verbal-semantic domain was exclusive to individuals with ASD, although it was not universal (see Section 2.7). It was reported that 30% of participants with ASD and 50% of the TD controls did not show a cross-domain bias. Loth et al. suggest several reasons for this finding and propose that local-global processing may be modulated by task demands, or could be independent constructs and not necessarily show a trade-off in performance.

2.12 Summary of the pervasiveness of central coherence and questions for the present study

To conclude, the consistency of central coherence within the same processing domain has mainly been studied in the visuo-spatial modality. This may be because experimental support for weak coherence in the ASD literature is more robust in this domain compared to the auditory/verbal modality. Findings have not been consistent and it has been considered whether the impact of general visuo-spatial abilities may drive associations between tasks, rather than processing style. Studies that have identified relative performance within an individual may resolve the problem of distinguishing style from ability: for example, determining whether the Wechsler Block Design subtest shows a peak relative to other Performance Scale subtests (Happé, 1994b), or applying Shah and Frith's (1993) methodology to determine an individual's benefit from segmentation, independent from ability to construct block designs. Integrative skills on copying and drawing tasks may again assess general visuo-spatial abilities and a coherent drawing style cannot necessarily be judged by the end product. The drawing process rather than product needs to be measured for a valid assessment of coherence.

It appears to be difficult to establish consistency in coherence across different processing domains. Establishing a reciprocal relationship between tasks that demand local processing and those that demand global processing has also been difficult to show empirically. Whether these are independent processes, or originate from one central driving force remains to be clarified. It is clear that assessment of local/global processing across processing levels/modalities is essential to address Frith's original conceptualisation of weak coherence, as a 'central force' to perceive and understand the world, pervasive in all levels of processing. There is also a need to operationalise performance that is considered as a deviation from normal, taking into account developmental level and gender.

The present study aims to address some of the outstanding questions emanating from the literature. To achieve this, data were collected from a large number of TD participants, on a range of coherence tasks designed to span visual and verbal/auditory modalities and low/high levels of processing. A balance was achieved by selecting or designing tasks that benefit from detailed processing in addition to tasks that place demands on global coherence. An attempt was also made to tap processing *style*, rather than ability by the use of open-ended tasks.

Although central coherence is a relatively new area of investigation within the autism literature, there is a large literature on the related field of local-global processing in typical development. The next chapter provides an overview of this diverse work, including the study of underlying mechanisms and developmental trajectories of local-global cognitive style. It is hoped that a review of the general literature will benefit our understanding of individual differences in central coherence in both TD and ASD populations.

Chapter 3. Central coherence in typical development

3.1 Introduction

Happé (1999) has suggested that central coherence may vary in the normal population and that people with autism may be at the extreme (weak) end of the normal continuum for this processing style. This fits with the notion long held by developmental psychologists that many dysfunctions in childhood could represent the extreme of a continuum rather than a categorically separate disorder (E. Taylor, 2001). It also highlights the mutually informative relationship that can exist between studies of typical and atypical development. Rutter (1989), for example, claimed "there can be no presupposition that normal and abnormal development do, or do not, involve the same mechanisms or do, or do not, share the same qualities" (p. 26). With regard to the present thesis, it is imperative to understand factors that influence the normal range of variation of central coherence, which might inform differences of clinical significance.

Although much debated, a drive towards gist and meaning, so-called "global coherence", is considered to be a universal feature of human information processing. Interest in the study of differences among individuals in this processing style has recently emerged, however, along with the search for the correlates and causes of such variability. The following chapter serves to introduce the field of local-global processing in typical development and review the work on individual differences in this processing style. As various conceptualisations of local-global processing exist in the literature, an exhaustive review of this vast area of research is beyond the scope of this thesis. Instead, an overview is provided of three relevant areas: (1) early theories that form the background to the notion of central coherence, (2) investigations into the developmental trajectory of local-global processing and underlying mechanisms involved, and (3) research on individual differences in local-global processing. The focus of this chapter is on general research findings and themes; more detailed information on individual tasks assessing central coherence can be found in Chapters 5 through 8.

3.1.1 A note on terminology

It is important to note the various terms that have been used in the literature to reflect different levels of stimuli structure: *part–whole*, *local–global*, *featural–wholistic*, *component–configural*. The same terms are used to suggest the level of information processing within an individual, as well as expressions such as, *bottom-up–top-down*, *analytical–wholistic*. These terms, although essentially alike, have come to represent slightly different conceptualisations and some authors suggest that they should not be used interchangeably.

Kimchi (1992), for example, made the useful distinction between local–global properties, as defined by "their position in the hierarchical structure of the stimulus" and featural–wholistic properties, defined "as a function of interrelations between the component parts of the stimulus" (p. 36). Wholistic properties, which Kimchi considered synonymous with configural properties, may be best captured by the Gestalt principles of perceptual grouping, such as closure, symmetry, and parallelism. Kimchi also suggests that correct use of the terms *top-down* and *bottom-up* processing should refer to conceptually driven and data-driven processing, respectively.

Problems of definition have been noted in experimental studies of local-global processing. Bhatt, Rovee-Collier, and Shyi (1994) draw attention to the fact that global and local are relative, not absolute, properties of stimuli. For example, a tree is a global form with respect to leaves and branches, but is also a local feature with respect to a forest. The ambiguity in what constitutes local or global level information and what an individual perceives as the local or global elements has led to questions over the validity of this research. To resolve this issue in part, Navon (1977) designed the compound or hierarchical figure stimulus, which comprises discrete and independent elements at the local and global level (e.g., a large square made up of small triangles). Such stimuli allow for greater experimental control through equating the global and local levels for factors such as complexity, familiarity, category, recognisability and processing load (Navon, 2003).

3.2 Historical perspectives of central coherence

The notion that human perception has a natural drive for coherence has a long history. The earliest conceptualisation of central coherence came from theories emanating from the Gestalt movement. With a focus on universals in information processing, Gestalt psychologists took the view that perceptual organisation is a top-down configuration of the whole, rather than a bottom-up sum of parts. Wertheimer (1924/1938) provided a classic example of this phenomenon in describing how melody is perceived:

> What I hear of each individual note, what I experience at each place in the melody is a *part* which is itself determined by the character of the whole. What is given me by the melody does not arise (through the agency of any auxiliary factor) as a *secondary* process from the sum of the pieces as such. Instead, what takes place in each single part already depends upon what the whole is. (p. 5).

The view that the quality of a part is determined by the whole gave rise to the main tenet of the Gestalt movement, 'the whole is more than the sum of the parts.' Wertheimer

(1938) maintained that the initial perception of spatial patterns were as unanalysed and undifferentiated wholes, dominated by grouping principles, and only later analysed into constituent parts. This perspective was in direct opposition to the structuralists (e.g., Titchener, 1909; Wundt, 1902) who held the view that the sensory whole must be built up from a conglomerate of elementary sensations. The Gestalt school of thought was also in contrast to the associationists (e.g., Hebb, 1949) who asserted that repeated viewings or connections between parts led to integration and formation of the conceptual whole.

The study of wholistic or global processing was not limited to visual perception; early experimental psychologists also examined semantic processing of visual and verbal material. Bartlett (1932), for example, referred to the 'drive for meaning' to describe the tendency of adults to recall the gist of information, rather than verbatim form. He also observed the tendency to introduce slight distortions during recall in order to make unfamiliar material more meaningful. In concluding his classic work on remembering images and stories, Bartlett stated:

> An individual does not normally take such a situation detail by detail and meticulously build up the whole. In all ordinary instances he has an overmastering tendency simply to get a general impression of the whole; and, on the basis of this, he constructs the probable detail. (p. 206).

As a modern extrapolation of the Gestaltists claims, Navon (1977) put forward the *global precedence hypothesis*, proposing that in adults, global properties are extracted from visual arrays and processed earlier than local properties. Visual processing is therefore suggested to be differentially sensitive to the hierarchical structure of stimuli. Presenting participants with hierarchical figures (in which larger letters are made up of smaller letters), Navon found reaction times for identifying global letters in adults were faster than those for local letters, an effect he called *global advantage*. Furthermore, when letters at the two levels were inconsistent, the global level interfered with responses made to the local level (*global interference*); in contrast, the local level had no effect on responses to global letters. The predominance of the global form led Navon to conclude, "global processing is a necessary stage of perception prior to more fine-grained analysis" (p. 371). It is interesting to note that global precedence may be limited to human beings. Fagot and Deruelle (1997) found an advantage of the local level with no interference from global level information in baboons even though these animals were able to process the visual stimuli globally. The local advantage in baboons was still apparent when the memory load of the task was removed and when the local elements were difficult to discriminate by being connected with lines or placed adjacent.

Navon's (1977) paradigm has been used extensively in the study of local-global processing across a range of psychological fields (see Navon, 2003, for a review). Despite the experimental control gained by the use of hierarchical figures, the perceptual relation between the global and local levels is sensitive to variations such as number and relative size of the local elements (see Kimchi, 1992, for a review). Kimchi claims that this finding weakens the underlying assumption that global form and the local elements map directly onto distinct perceptual units, differing only in level of globality. She proposes that the study of wholistic properties of stimuli, that is, properties that depend on the configuration of component parts, may be more informative when testing the precedence effects in perceptual processing.

3.3 The developmental course of local-global processing

One of the most enduring questions in developmental psychology is whether the primacy of processing parts versus wholes systematically changes with development. The study of part-whole perception in typical development has received great scrutiny across a number of diverse research fields such as featural and configural visual-perceptual processing (e.g., Kimchi, 1992), visuo-spatial construction (e.g., Akshoomoff & Stiles, 1995a), and coherence and comprehension in language (e.g., Gernsbacher, 1993). Appendix A presents a summary of published experimental studies that have examined the development of local-global processing with examples selected from infancy, childhood, adolescence and adulthood populations. An overview of the main findings from each field is provided below.

3.3.1 Local-global processing during infancy

Gestalt psychologists held the view that the rules that govern perceptual organisation were innate in the visual system (Koffka, 1935). However, constructivists (e.g., Hebb, 1949; Piaget, 1952) suggested that perception is built up from the separate perception of lines and angles, which become organised or constructed with development. There is growing empirical evidence of a developmental change in the perception of parts and wholes in the first few months of life (see Colombo, 2001, for a review). Several studies report a developmental sequence from local to global processing during infancy, with some suggesting that global processing may predominate in infants as young as 3 months (Bhatt et al., 1994; Ghim & Eimas, 1988; Van Giffen & Haith, 1984; Younger & Cohen, 1986). Cohen and Younger (1984) suggest that a shift in the perception of form occurs sometime after 6 weeks of age. Prior to this age infants can remember (i.e., habituate to) the specific orientations of line segments, but not the angular relations that line segments can make. Consistent with the constructivist view, Cohen and Younger propose, "only later through experience and/or development are [infants] able to integrate those line segments into entire angles and shapes" (p. 46). However, under certain experimental conditions newborn infants have been shown be sensitive to the emergent properties of line elements rather than component lines. Slater, Mattock, Brown, and Bremner (1991) demonstrated that repeated exposure of the same angle in different orientations led to habituation to that angle in newborns. Slater et al. propose the possibility that "form perception may not be dependent upon a lengthy period of learning and/or maturation for its development" (p. 405).

3.3.2 Local-global processing during childhood and adolescence

It is well documented that perceptual processing undergoes much change during childhood, being fundamentally different from and more limited than that of adults. However there appears to be little consensus on the exact nature of this change. Early developmental studies that questioned whether parts or whole dominate children's perception generally provided mixed results. Some studies asserted that young children are best described as 'piecemeal' processors, attending only to the parts of a configuration (e.g., Carey & Diamond, 1977; Corah & Gospodinoff, 1966; Elkind, Koegler, & Go, 1964); other studies claim that young children are "wholistic" processors, perceiving patterns as undifferentiated wholes, without awareness of constituent parts (e.g., Gibson, 1969; Meili-Dworetzki, 1956). This latter view was held by early developmental theorists (e.g., H. Werner, 1957) who proposed that perception progresses from a state in which parts are not distinguished from the whole, to the state in which parts and wholes are differentiated. Werner provides an example of this developmental change in his discussion of children's descriptions of Rorschach inkblot stimuli:

> ... with age there is a decrease of the undifferentiated diffuse whole and detail responses along with an increase of the highly articulated, well-integrated whole and detail responses. There is further an interesting shift from the early whole responses toward small detail responses between the ages of about six and eight; later on there is a decline in favour of the integrated whole responses. (p. 141)

Dukette and Stiles (1996) point out that the historical emphasis on parts *versus* wholes may have produced a false dichotomy. Depending on task and stimulus conditions, young children have demonstrated an ability to attend to either the parts, to the whole pattern, or both (Dukette & Stiles, 1996, 2001; Tada & Stiles-Davis, 1989). Rather than an isolated ability to perceive parts or wholes during different development stages, Aslin and Smith (1988) stress that what changes is the "quality of the representation and the available operations that can be performed on it" (p. 458). Along a similar line, Tada and Stiles (1996) state the question posed in more recent developmental studies is "not whether children are wholistic or analytic processors of information, but what is the nature of this process and how does it change with development" (p. 952). The specific developmental trajectory of local-global processing is likely to be affected by the nature and difficulty of the task. Developmental changes in cognitive skills involved in many local-global tasks, such as identification, classification, and discrimination, also need to be taken into account.

The emphasis across the plethora of developmental studies appears to be on the proficiency with which information (visual or verbal) can be integrated into a coherent whole with age, for example to extract gist (Brainerd & Gordon, 1994; Reyna & Kiernan, 1994), demonstrate global precedence (Kimchi, 1990; Kramer, Ellenberg, Leonard, & Share, 1996), or to show coherence in copying tasks (Akshoomoff & Stiles, 1995a; Tada & Stiles, 1996). Tasks that require detail-focused processing and an ability to ignore gestalt principles (e.g., EFT, Block Design) also show greater proficiency with age (e.g., Enns & Girgus, 1985; Pennings, 1988; Witkin, Oltman, Raskin, & Karp, 1971). It appears, then, that the reciprocal processes of integration and segmentation show developmental progression. Studies have shown that pre-school children can accurately analyse and integrate elements of a pattern in tasks of drawing, block construction, and perceptual judgment (e.g., Dukette & Stiles, 2001; Kimchi, 1990; Stiles, Delis, & Tada, 1991; Tada & Stiles, 1996; Tada & Stiles-Davis, 1989). However the processes by which they achieve this appear to differ from that in older children. Pre-school children tend to specify many small parts and overly segment a form; older children in contrast, are more sensitive to larger components and can synthesise these with more complex spatial relations.

It is difficult to make general conclusions from the range of developmental studies that exist, however in a comprehensive review of published studies on local-global processing in children, Aslin and Smith (1988) surmise that the developmental trend in part versus whole perception appears to be curvilinear: "parts dominant perception during infancy, parts become organised into wholes during early childhood as parts become relatively inaccessible, and finally parts reemerge as analysable constituents of whole objects" (p. 458). Interestingly, Mondloch, Geldart, Maurer, and de Schonen (2003) report that developmental changes in processing hierarchical figures continue into adolescence, possibly reflecting ongoing development in brain structures such as the frontal lobe during this time (Giedd, Blumenthal, Jeffries, Castellanos et al., 1999).

Developmental studies assume the progression of local-global processing to be the result of intrinsic maturation and a question rarely considered is the possible contribution of environmental factors. One example is the processes involved in learning to read. Initially the child is instructed to attend to individual components of words, and only once the child becomes a more efficient and confident reader, is emphasis given to the context to aid word decoding. It is as yet unclear whether certain teaching methodologies impact on how children process local versus global information.

3.3.3 The effect of aging on local-global processing

Interest in the cognitive changes associated with normal aging has prompted research into local-global processing in later life. Although only few studies exist, Bruyer, Scailquin, and Samson (2003) provide a useful review of the effects of aging on global precedence. Some studies report enhanced detailed processing in the elderly, with a local advantage on tasks using hierarchical figures (Polster & Rapcsak, 1994; Stark & Coslett, 1993). This may be associated with a difficulty in integrating information at a global level, an ability that has been shown to decline with age (Ludwig, 1982). Some studies, however, report no change with age on the global precedence effect commonly found in young adults (Bruyer & Scailquin, 2000; Bruyer et al., 2003), or even greater inference effects of global information on local processing with age (Roux & Ceccaldi, 2001). Bruyer and Scailquin found that subgroups existed in their elderly sample, differing in sensitivity to these interference effects, but not in age or crystallised intelligence. They further conclude that interference effects did not appear to result from a general deficit in the inhibitory mechanisms that are suggested to decline with age (Hasher, Stoltzfus, Zacks, & Rypma, 1991). Although the research findings to date are inconclusive, Bruyer et al. propose that studies of developmental change across the life span may help elucidate the underlying mechanisms involved in local and global processing.

3.4 Mechanisms underlying local-global processing

Along with characterising the developmental trajectory of local-global processing, there has been a long interest in defining the underlying mechanisms of this process. The following provides a review of studies that have investigated possible mechanisms of coherence from a developmental perspective.

3.4.1 The role of perceptual regulations

Elkind and colleagues (Elkind, Anagnostopoulou, & Malone, 1970; Elkind et al., 1964) observed the developmental progression of children's part-whole descriptions of objects, when both levels were meaningful (e.g., a picture of a face made of fruit). Young children (4 to 6 years) tended to identify local parts or the global whole only and older children (8 years and above) would reliably report both the parts and the whole spontaneously. It was proposed that with development comes mental flexibility, which enables an individual to perceive two or more entities within the same figure (cf. perception of ambiguous figures; Rock, Gopnik, & Hall, 1994). Elkind and colleagues interpreted this perceptual shift with respect to Piaget's developmental theory of perception, and in particular to *decentering* theory: the ability to examine the perceptual field from several perspectives, which is suggested to develop in middle childhood (Piaget, 1969). Piaget also proposed that *perceptual regulations* develop with age, which allow perception to be governed by laws similar to logic. These regulations are suggested to replace or integrate with the more elementary Gestalt organisations in perception. Elkind and colleagues suggested that the development of both processes allows a child to perceive both local and global aspects of a figure, and perform the logical calculations necessary to appreciate that an object can belong to two classes at once.

3.4.2 Executive control and shifting between local and global levels

In accord with Elkind's description of the mental flexibility required in local-global processing (Elkind et al., 1964), recent research has also considered the role of executive control when attending to local-global information. Age-related changes, on this account are not due to a tendency to perceive either the larger global form or the local details, but rather due to increasing ability to adapt attentional strategies to the demands of the task (Enns & Girgus, 1985; Tada & Stiles, 1996). Such strategies may include the ability to shift from a feature to the whole, switch between processes of analysis and synthesis, and to ignore an irrelevant level. It is well documented that many facets of executive functioning, such as the efficiency of working memory capacity, planning and problem-solving, develop with age (De Luca et al., 2003), as does the ability to ignore distractors (Goldberg, Maurer, & Lewis, 2001; Ridderinkhof & van der Molen, 1995). The role of executive skill in local-global processing, and the possible interaction between these processes across development, warrants further examination.

3.4.3 Perceptual versus conceptual level of processing

Several studies have proposed that global advantage is mediated by sensory mechanisms (e.g., Navon & Norman, 1983; Shulman, Sullivan, Gish, & Sakoda, 1986). Navon's (1977) original interpretation of global precedence was that it reflected a priority of global properties at an early pre-attentive or perceptual stage of processing. Although the finding that baboons show local advantage despite a largely human-like sensory system questions whether global precedence is based purely on sensory processes (Fagot & Deruelle, 1997). There is further contention concerning whether global advantage may arise at a post-perceptual stage or throughout several stages of processing (see Kimchi, 1992, for a review). Bruyer et al. (2003) recently addressed this question by manipulating hierarchical figure stimuli of the standard Navon task. Local and global letter stimuli were randomly presented in lowercase and uppercase in order to force a grapheme representation of the target letter, considered to require late perceptual processing. Global precedence effects were found in young and old adults using both predictable and unpredictable targets, providing evidence of priority of global processing at different stages of processing. To attain a comprehensive assessment of local and global processing it is clear that measurement should be made across levels of processing ranging from perceptual-sensory through to high-level conceptual.

3.4.4 Independence of local and global processing

The notion that local and global processing rely on different mechanisms has emerged from several different lines of evidence. Neuropsychological studies of brain-injured patients, for example, have broadly-speaking shown that unilateral lesions of the left cerebral hemisphere impair the processing of local level information, whereas lesions of the right hemisphere disrupt the processing of the global level (for a review, see L. C. Robertson & Lamb, 1991). Functional neuroimaging research also broadly supports the role of the left hemisphere in processing local elements, and of the right hemisphere in processing the global form in TD individuals (e.g., Heinze, Hinrichs, Scholz, Burchert, & Mangun, 1998; but stimulus type may play a role, see Fink et al., 1997; Fink, Marshall, Halligan, & Dolan, 1999; also see Chapter 2, Section 2.6.2).

Recent work into the developmental trajectory of local-global processing has also led researchers to the premise that independent mechanisms may exist in processing local and global level information (Burack, Enns, Iarocci, & Randolph, 2000; Dukette & Stiles, 1996, 2001; Porporino, Shore, Iarocci, & Burack, 2004) and a brief review of these studies follows.

The study of hierarchical figures analysis in children led Dukette and Stiles (1996; 2001) to propose that the reciprocal processes of integration and segmentation differentially change with age. In a similarity judgment task, younger children (4 to 6 years) were more likely to make judgments based on local details than global similarities, specifically when the density of the hierarchical figures was sparse (Dukette & Stiles, 1996). When required to draw hierarchical figures from memory, younger children were better able to draw local elements than global form (Dukette & Stiles, 2001). Again, this was particularly evident with sparse stimuli, which demanded greater integration than segmentation processes. Dukette and Stiles (2001) interpret these findings as evidence that "integration and segmentation processes of spatial analysis develop along somewhat different time courses, with segmentation abilities becoming more elaborated earlier than integration skills" (p. 247).

The developmental studies of Burack and colleagues were initiated by the work of Enns and Kingstone (1995) who applied the visual search paradigm to the study of local-global processing in adults. Participants were instructed to search for a triangle feature that could appear as a local element (i.e., a square made up of local triangles) or a global element (i.e., a triangle made up of local squares) amongst a variable number of distractor stimuli (global squares made from local squares). The size and density of local elements was varied in order to manipulate local and global precedence effects (i.e., small size and dense spacing favoured global detection). Under various conditions, search times for global targets were found to increase with the number of distractors present. In contrast, search times for local targets did not vary as a function of number of distractors. Enns and Kingstone hypothesised that different processes underlie global and local detection: parallel or simple sensory processes may be in operation during local element detection, while serial or more sophisticated processes may be required for global level detection.

Perception of the global structure in a hierarchical figure during visual search may require a grouping operation, over and above the perception of the local elements. To test whether this could explain the difference between local and global processing, Enns and Kingstone (1995) designed stimuli that required grouping at both the local and global level. Distractor stimuli consisted of two pairs of vertically aligned dots; local targets consisted of a feature change (i.e., the oblique alignment of dots within each pair), while global targets consisted of a configural change (i.e., the oblique alignment of dot pairs). As with the traditional hierarchical stimuli, the increase of distractors added to the attentional demands of global target detection, but not local target detection. To increase the difficulty of local target "grouping", dot colour was contrasted within pairs. This rendered an increase in attentional demands in local target detection, but not in global target detection as the contrast in dot colour more easily defined the global stimulus. The authors concluded, "the grouping of noncontiguous elements in compound stimuli is an attention-demanding operation, regardless of whether that operation is required at the local or the global level" (p. 289).

Burack et al. (2000) applied this experimental design to the study of developmental differences in processing of local and global features during visual search. Using the same dot stimuli (two pairs of vertically aligned dots), Burack et al. defined the manipulation of spatial distance among elements involving an oblique shift within dot pairs as *short-range* grouping and an oblique shift between dot pairs as *long-range* grouping. Age improvements in visual search (as defined by the impact of distractors on search times) were found for long-range targets, whereas search for short-range targets was constant across age (6 to 22 years). This finding held even when the perceptual access to the local or global level of stimuli was made difficult by contrasting the dot colour, within or between pairs. In discussing the implications of their findings, the authors suggest that developmental effects may operate on sensitivity to global level structure in a pattern, independently from local level perception. Furthermore, adult-like efficiency may be attained earlier for local than for global processing.

Porporino et al. (2004) also assert that the developmental trajectories for local and global processing involve independent mechanisms. This was concluded from a study of selective attention in which participants responded to target shapes (square and diamond) that could appear at either the local or global level of a hierarchical figure, while ignoring the interfering effects of distractor hierarchical figures positioned either side of the target. Global advantage was found across all age groups (6 years to adult) when no distractors were present, and when hierarchical levels of the target and distractor were identical. The youngest participants (6 to 8 years) were affected by the presence of neutral distractors (as shown by increased reaction times) but, interestingly, only when responding to global targets and not local targets. It was suggested that this provided evidence for separate developmental trajectories for local and global level processes with efficient local processing attained early and global processing continuing to improve at least until 8 years of age. Alternatively, the findings may be interpreted in terms of developmental effects in the ability to control attentional focus. Pastò and Burack (1997) for example, have previously reported that ability to adjust the range of attentional focus increases after 9 years of age. Attending to the global level of the target may be more vulnerable to interference effects of distractors in young children, more so than when attending to local level information.

Evidence for independent mechanisms for local and global processing has also been found in the verbal-semantic modality. Memory studies by Reyna and Kiernan (1994) and Brainerd and Gordon (1994) both concluded that memory for gist was not dependent on verbatim recall, particularly when participants were directed to remember information word-for-word. This finding dispels the assumption held in language research that verbatim representations evolve into gist. Instead, memory appears to be both constructive and hold the ability to preserve details (Reyna & Kiernan, 1994). Consistent with the finding of Burack and colleagues that local processing may be attained earlier than global processing (Burack et al., 2000; Porporino et al., 2004), Brainerd and Gordon reported that gist was relatively higher than verbatim recall in 8-year-olds, with the opposite pattern of verbatim better than gist recall in 5-year-olds. Furthermore, the two age groups did not differ in absolute level of verbatim recall.

In sum, there is growing evidence to suggest that local and global processing are independent constructs, with local processing being attained earlier in development than global processing. Attempts to define the underlying mechanisms of local-global processing have also led to the question of whether this processing style is independent of executive and attentional control.

3.5 Individual differences in central coherence

Cognitive and developmental psychologists have traditionally focused on universals in local-global processing, although interest in individual variation can also be found. Early developmental studies, for example, acknowledged the fact that not all children followed the universal progression in local-global development. Elkind et al. (1964) reported a developmental shift in appreciating both parts and wholes at approximately 8 years, although individual variation was noted with 30% of 8-year-olds and 20% of 9-year-olds reporting local parts only. Elkind et al. (1970) also highlighted individual variability in the development of perceptual regulations. They reported that a small proportion of children (4 from 20) who were well able to perceive both parts and wholes in drawings, surprisingly failed a test of mathematical logic that tested understanding of elements within a set, such as appreciating that objects can belong to two classes at once. Elkind et al. (1970) acknowledged, "while part-whole perception may be mediated by perceptual regulations for the majority of children, it may not be so mediated in all children" (pp. 396-397).

Individual differences have also been examined in infant research. In particular, individual variations have been found in the amount of time infants tend to fixate on visual stimuli, which may reflect differences in visual information processing. *Short-duration* infants are suggested to take in information in a global to local sequence, whereas *long-duration* infants may process information immediately at a local level and perform a feature-by-feature analysis (Colombo, Mitchell, Coldren, & Freeseman, 1991; Freeseman, Colombo, & Coldren, 1993; Stoecker, Colombo, Frick, & Allen, 1998). As a follow-up to this finding, Frick, Colombo, and Allen (2000) examined whether individual differences in look duration were related to differences in processing hierarchical patterns. Short-duration infants showed global precedence effects after viewing a hierarchical pattern for 20 seconds (i.e., displayed a novelty preference to global, but not local properties of the familiarised pattern). Global precedence effects were also present in the long-duration infants, but only after they were familiarised with the hierarchical pattern for 30 seconds. Stoecker et al. discuss whether prolonged looking could be related to an inability to disengage or "inhibit" looking at features. In their study long-duration infants were found to show no benefit from the symmetry of stimuli (typically processed faster than asymmetrical stimuli), possibly perseverating on a part of the stimulus in preference to the whole form. Fixation duration in infancy has also shown an association with later general cognitive abilities, with longer durations associated with lower intellect (Colombo & Mitchell, 1990).

The investigation into individual differences was informative in the study by Dukette and Stiles (1996) on developmental effects on forced-choice matching using hierarchical figures. Individuals were categorised as *local* or *global* responders if the number of matches made to either level significantly differed from chance, otherwise as *ambivalent* if neither level was significantly biased. The majority of children demonstrated a significant response bias, mostly towards global level matches: 8 out of 12 4-year-olds, and 10 out of 12 6-year-olds demonstrated a significant global bias. When the density of local elements was reduced, the majority of children demonstrated a response bias (i.e., 10 out of 12 4- and 6-year-olds), but a shift towards local level matching was more prevalent in the youngest age group (i.e., 8 of 10 4-year-olds showed a local bias, 9 of 10 6-year-olds showed a global bias). The individual analysis therefore provided important information concerning the nature of the shift in response patterns; younger children were not simply more ambivalent about their choices, but demonstrated a change in response bias from the global to the local level.

3.5.1 The field dependence-independence distinction

A key movement in individual difference research has been the work on the cognitive style of 'field dependence-independence' by Witkin and colleagues (Witkin & Goodenough, 1981; Witkin et al., 1971). Using tests in which participants were asked to align a rod to a vertical position while ignoring a surrounding tilted frame, Witkin found that individuals scored on a continuum with two extreme groups: those who could perform the task easily, and those who encountered great difficulty. Individuals who showed proficiency on the task were labelled as *field independent*, referring to a tendency to perceive objects as discrete units, distinct from the field as a whole. In contrast, individuals labelled as *field dependent* were described as having a cognitive style characterised by perception that is strongly dominated by the overall organisation of the surrounding field, with the parts of the field experienced as "fused" (Witkin et al., 1971). Witkin observed the consistent tendency among individuals to be field-dependent or field-independent across tasks, most notably the EFT. Gestalt psychologists introduced EFT tasks as they proposed effort was required to resist the tendency to perceive the strong gestalt of the complex form (Koffka, 1935). Individuals categorised as having a field independent cognitive style are therefore defined by good EFT performance.

Although a similar construct to central coherence, it is incorrect to equate weak central coherence with field independence (Happé & Frith, 2006). While field independent people are thought to succeed on the EFT through an ability to overcome the gestalt, individuals with weak central coherence are suggested not to spontaneously attend to the gestalt, and instead perceive the local elements in the initial stage. Happé and Frith propose that this distinction may be tested empirically through analysis of error-types made on the Block Design subtest: individuals with weak coherence are predicted to make construction errors that preserve details but violate the configuration of the design; in contrast, field independent individuals are predicted to make errors that preserve the configuration but violate the details. Thus the default processing style in field independent individuals (similar to field dependent individuals) is global processing, whereas in individuals with weak central coherence the default is local processing.

The theory of field independence/dependence petered out because of the emerging evidence that cognitive *ability* and not style was primarily being tapped by measures such as the EFT (J. A. Richardson & Turner, 2000; Vernon, 1972). Field independence is often associated with higher intelligence, however it remains to be seen whether weak central coherence can be verified as a cognitive style, independent from ability.

3.5.2 Intelligence and local-global processing

As mentioned in the previous section, the role of intelligence on local-global processing needs to be acknowledged in the study of individual differences in processing style. Tasks requiring detail-focused processing that show greater proficiency with age (e.g., EFT, Block Design) often show an association with general intelligence (see Chapter 5). It is of interest to know whether tasks intended to tap processing *style* or bias, are also dependent on intellectual abilities. Dulaney, Marks, and Devine (1994) investigated this by comparing adults of low intelligence and TD individuals on tasks utilising hierarchical figures. Although low-functioning individuals made a greater number of errors relative to the TD sample, global precedence effects were found in both groups on tasks where attention was directed to either the global or local level. With repeated practice, both groups improved in identifying local level information, although improvement at the global level was only demonstrated in TD individuals. Global precedence effects were still apparent in both groups after repeated practice. Dulaney et al. conclude that their findings are consistent with studies that have shown greater intelligence-related differences in tasks that are more effortful and attention demanding than in less effortful tasks (e.g., Fox & Oross, 1990; Soraci & Carlin, 1992).

3.5.3 *Gender effects in local-global processing*

In addition to individual effects of age and IQ, gender effects have also been found in local-global processing. On visual coherence measures where detail-focus is advantageous (e.g., EFT, Block Design), gender differences are often reported with males surpassing females in speed and accuracy (see Chapter 5 for a review). Tasks that tap global or local processing preference, rather than ability, have produced mixed results however. Kramer et al. (1996) found that males (4 to 12 years) tended to match hierarchical geometric figures based on global form, significantly more so than females. Using the same similarity judgment paradigm with letter stimuli, Dukette and Stiles (1996) report the opposite effect, with females making significantly more judgments based on the global letter than males (4 to 6 years, and adults). Interestingly the different stimuli used may inform the opposing findings, with females being particularly biased to make similarity judgments on the basis of the larger, global letter, rather than geometric forms. Dukette and Stiles also used geometric shapes to ensure the 4- and 6-year-old children in their sample were as familiar with the stimuli as older participants, and found no effect of gender.

Kramer et al. (1996) interpret their findings as being consistent with developmental models that propose an early left-hemisphere (e.g., verbal skills) advantage for females and a right-hemisphere (e.g., visuo-spatial skills) advantage for males. This is in keeping with the historic view that male superiority on tasks requiring visuo-spatial abilities are "among the most persistent of individual differences in all the abilities literature" (McGee, 1979, p. 41). Current research shows that while males tend to outperform females on visuo-spatial tasks, effect sizes tend to be small, suggesting that the overlap in the distribution of male and female scores dwarfs the difference between them (Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003). However, if existent, a male superiority on visuo-spatial tasks may be shown as a bias toward local processing (e.g., superior segmentation skills) or superior global integrative abilities depending on the task demands. Typically the processing requirement (local/global) and modality (visual/verbal) are confounded in studies. To obtain a complete picture of male-female differences in detail-focused processing, assessments are required in auditory and verbal domains, independent of general ability differences. A comprehensive assessment of processing style across different modalities is one aim of the present thesis, and will enable a thorough examination of gender differences in local-global style.

3.5.4 *The pervasiveness of local-global processing style in typical development*

A central question of this thesis is whether stable individual differences exist in central coherence across processing modalities. Only a few investigations have assessed the pervasiveness of cross-domain processing bias in TD populations. Gernsbacher and colleagues, for example, have concluded that the same cognitive mechanism appears to underlie reading comprehension of written text (presumably involving coherence at a semantic level) and comprehension of nonverbal picture stories (Gernsbacher et al., 1990). Gernsbacher demonstrated that ability to suppress irrelevant material appears to be key in comprehension, whether the material is linguistic, non-linguistic, or cross-modality (Gernsbacher & Faust, 1991). This proposal of less efficient suppression mechanisms during sentence processing has direct relevance to studies of central coherence in autism (e.g., weak coherence on the Homograph Reading Task, Happé, 1997; and Sentence Completion Task, Happé & Booth, 2006; see also Chapter 8).

Stable individual differences in processing style have been captured on tasks within the same processing domain. Kramer, Kaplan, Share, and Huckeba (1999) for example, observed that children (aged 6 to 12 years) who made configural errors on block design tasks (i.e., constructions that break the overall configuration of the pattern but preserve detail) showed a tendency to match by local element on a similarity judgment task using hierarchical figures. As mentioned above, Witkin (Witkin & Goodenough, 1981; Witkin et al., 1971) observed stable individual differences in the dimensions of field dependence-independence across different tasks, but did not examine this processing style across different modalities. A recent study by Pellicano, Maybery et al. (2005; described in Chapter 2, Section 2.11.1) examined whether stable individual differences in processing style in the visuo-spatial domain could be found in a group of preschool children. They concluded that central coherence was not a unitary construct; good visual integration was associated with good visual analysis, and did not show a trade-off in local-global processing. It is clear that further study needs to address the issue of whether stable individual differences in local-global processing exist across different processing levels and modalities.

3.6 The present thesis

The study of local-global processing in typical development has a long and extensive history. Many aspects of the work conducted in this area can be applied to the study of weak central coherence in ASD. It is imperative to understand whether the unusual detail-focused processing style observed in autism is at the extreme end of a normal continuum, or whether it is somehow qualitatively different.

From the review presented in this chapter it may be concluded that very few studies have examined whether stable individual variation in processing style is present across various modalities and levels, that is not simply a reflection of ability. For example, does a local perceptual style in the visual modality (e.g., analysing hierarchical figures) predict detail-focus in verbal coherence (e.g., reading comprehension).

Local and global processing has been found to develop with age, possibly at different rates and/or independently. The conceptualisation of coherence as a trade-off between local versus global processing (e.g., when viewing hierarchical figures), may limit our understanding of the underlying processes involved. Task design should be sensitive to measuring the demands of both local and global processing, and consider these as independent constructs. The role of executive control, such as the ability to switch between levels, also needs to be acknowledged within task design. In order to clarify gender differences in coherence, tasks also need to span visual and verbal domains.

The current thesis provided a thorough examination of individual differences in central coherence, in order to take these central issues into account. This was achieved by the development of a task battery designed to assess individual differences in coherence at different levels of processing (low-level, high-level) and modality (visual, verbal, auditory). The task battery was administered to a large sample of TD individuals chosen to span a wide age range (8 to 25 years). Effects of age, intellectual ability, and gender on processing style in typical development were examined. Performance on the task battery was also compared between individuals with ASD and age- and IQ-matched control participants, in order to determine whether weak coherence is universal in individuals with autism, and correlated across processing levels/modalities.

In sum, the aim of the thesis is to address the hypotheses that (1) pervasive individual differences can be found in coherence in TD children and adults, across levels and domains of processing; and (2) individuals with ASD have a pervasive, unique and universal deficit in central coherence, which is manifest as a bias towards weak coherence. The general methodology used in the study of individual differences in central coherence in TD and ASD is presented in the following chapter (Chapter 4). Details of individual tasks and experimental findings are presented by processing-level/modality in Chapters 5 through to 8.

Chapter 4. General Methodology

4.1 Introduction

The aim of this chapter is to describe the general methodology of the study on central coherence, to outline the research design, and describe the data collection and methods of analyses. Specific details of the methodology, analysis, and results for each task included in the coherence task battery will be described in Chapters 5 through 8.

4.2 Participants

4.2.1 Typically developing participants

In total, 205 typically developing (TD) participants aged from 8 to 25 years took part in the study of normal variation in central coherence.

School-aged TD participants were recruited from one secondary school and two primary schools in South London, and two secondary schools in South East England. Permission to undertake the study in schools was initially obtained from the respective head teacher. Class teachers were asked to randomly select children and letters were sent home to parents and guardians detailing the nature of the study and what their child's participation would involve.

The majority of TD adults were recruited while the researcher was based at the University of Auckland, New Zealand. Advertisements were placed in various locations (student job centre, public library, youth hostel), inviting people aged between 17 and 25 years to take part in a study of problem-solving styles. As the majority of participants recruited were University students (28 from 31) and had high full-scale IQ (FIQ) estimates (M = 115, SD = 12), further recruitment of lower ability participants was necessary to obtain a more heterogeneous sample. Advertisements were placed on hospital notice boards, shop-windows and a youth club in South London, which resulted in recruitment of nine further participants (FIQ M = 95, SD = 14).

All TD participants were administered a short-form of either the Wechsler Intelligence Scale for Children (WISC III; Wechsler, 1992) or the Wechsler Adult Intelligence Scale (WAIS III; Wechsler, 1997a) according to age. The short-form was based on four subtests: Information and Picture Completion, selected due to their short administration and scoring time (Kaufman, Ishikuma, & Kaufman-Packer, 1991); and Block Design and Vocabulary, selected due to their high reliability and validity estimates (Cyr & Brooker, 1984). The IQ estimate calculated from this combination is reported to have high reliability (Sattler, 1992). Tellegen and Briggs's Formula 4 (1967, p. 504) was used to convert the sum of scaled scores for the WISC-III or WAIS-III subtest combinations into FIQ, verbal (VIQ) and performance IQ (PIQ) equivalents. This formula takes into account the correlations between the subtests and is a widely used method of obtaining estimates of IQ (e.g., Jeyakumar, Warriner, Raval, & Ahmad, 2004). Participants were required to have a minimum FIQ estimate of 70 for inclusion in the study. This excluded one participant from the original sample. Further exclusions were made if participants had any clinically significant impairment or diagnosis, or family history of ASD (no exclusions made on this basis).

Participants were banded into four age groups (8-10 years, 11-13 years, 14-16 years, 17-25 years) to allow examination of developmental trends on the coherence measures. Full participant characteristics by age group are shown in Table 4.1. Age groups were matched on FIQ and VIQ, although a PIQ advantage was observed in the oldest age group ($F_{(3,200)}$= 4.35, p = .005). As the Block Design subtest is regarded as a standard measure of coherence in the autism literature (Shah & Frith, 1993), a second FIQ score was calculated which excluded Block Design (FIQ-BD, Table 4.1) in order to eliminate any confound with other coherence measures in later analyses. No age group differences were found in FIQ-BD.

Table 4.1 TD sample characteristics by age group: Mean (*SD*).

Age group (years)	N	FIQ	VIQ	PIQ[a]	FIQ-BD
8-10	54	107.2 (14.7)	110.6 (17.1)	101.3 (12.5)	108.9 (14.4)
11-13	43	106.0 (15.5)	106.7 (16.2)	103.1 (14.0)	106.7 (15.2)
14-16	51	104.6 (13.0)	108.0 (16.1)	99.8 (10.6)	104.7 (13.9)
17-25	56	111.4 (14.2)	111.6 (13.9)	108.4 (15.5)	108.9 (13.4)

[a]17-25 > 8-10, 14-16, p < .01

Participant characteristics divided by age group and gender, are presented in Table 4.2. Chi-square tests showed that the distribution of male and female participants across the four age groups was statistically equivalent (χ^2 = 5.95, p = .11). A two-way ANOVA (with age group and gender as independent variables) found no main effect of gender on age ($F_{(1,196)}$= 3.08, p = .08). Male and female participants were also matched in age within each age group with no age group by gender interaction found ($F_{(3,196)}$= 1.71, p = .17).

A series of two-way ANOVAs on each IQ measure revealed a moderate main effect of gender on FIQ ($F_{(1,196)}$= 3.84, p = .052) and significant interactions between age group and gender on FIQ ($F_{(3,196)}$ = 2.74, p = .05), VIQ ($F_{(3,196)}$= 4.59, p = .004) and FIQ-BD ($F_{(3,196)}$= 3.37, p = .02). Overall, males had higher FIQ than females, which approached statistical significance ($t_{(175.9)} = 1.85$, $p = .07$, unequal variances). By age group, a male superiority in FIQ, VIQ and FIQ-BD was apparent in the youngest (all $t_{(52)} > 1.98$, $p < .05$) and oldest groups (all $t_{(54)} > 2.04$, $p < .02$). In the 14-16 year group the gender effect was reversed with females performing higher on VIQ than males ($t_{(46.7)} = 2.02$, $p = .05$).

Table 4.2 TD sample characteristics by age group and gender: Mean (*SD*).

Age group (years)		N	Age	FIQ[a]	VIQ[b]	PIQ[c]	FIQ-BD[d]
8-10	*male*	19	9.7 (0.8)	112.4 (16.3)	118.9 (18.6)	101.7 (12.5)	114.5 (15.8)
	female	35	9.7 (0.6)	104.4 (13.1)	106.2 (14.7)	101.1 (12.6)	105.9 (12.9)
11-13	*male*	21	12.3 (0.9)	107.0 (16.4)	107.2 (16.4)	104.4 (15.3)	107.1 (14.9)
	female	22	12.3 (1.1)	104.9 (15.0)	106.3 (16.5)	101.8 (12.9)	106.4 (15.8)
14-16	*male*	30	15.4 (0.8)	102.8 (15.3)	104.7 (18.6)	100.1 (11.4)	102.3 (16.0)
	female	21	15.7 (0.8)	107.2 (8.6)	112.9 (10.1)	99.3 (9.6)	108.1 (9.7)
17-25	*male*	27	19.9 (2.0)	116.6 (15.3)	116.3 (14.5)	112.7 (15.5)	113.7 (14.6)
	female	29	21.0 (2.6)	106.6 (11.3)	107.1 (11.9)	104.4 (14.6)	104.4 (10.5)
All	*male*	97	14.9 (4.0)	109.4 (16.4)	111.2 (17.8)	104.9 (14.4)	108.9 (16.0)
	female	107	14.5 (4.8)	105.7 (12.2)	107.8 (13.6)	101.8 (12.7)	106.1 (12.3)

[a]Male 14-16 < 17-25, $p = .007$
[b]Male 14-16 < 8-10, $p = .03$
[c]Male 8-10, 14-16 < 17-25, $p < .05$
[d]Male 14-16 < 8-10, $p = .04$

Effects of age group on IQ were examined for male and female participants separately. While significant age group effects were found for males on all measures (FIQ $F_{(3,93)} = 4.03$, $p = .01$; VIQ $F_{(3,93)} = 3.93$, $p = .01$; PIQ $F_{(3,93)} = 4.46$, $p = .006$; FIQ-BD $F_{(3,93)} = 3.67$, $p = .02$), none was found for female participants (all $F_{(3,103)} < 1.26$, $p > .28$). Post hoc analyses (Tukey's HSD) revealed a general inferiority in IQ in males from the 14-16 year group compared to males in the 8-10 and 17-25 year groups (see Table 4.2).

4.2.2 *ASD and Control participants*

A total of 32 males aged from 9 to 21 years with diagnosed autism spectrum disorder (25 Asperger's Disorder, 7 Autistic Disorder) took part in the study of central coherence in ASD. In each case it was confirmed that a psychiatrist or paediatrician had made the diagnosis according to established criteria in the DSM-IV (American Psychiatric Association, 1994). Social and Communication Questionnaire (SCQ; Berument, Rutter, Lord, Pickles, & Bailey, 1999) data were available for 24 participants. One participant (with an Autistic Disorder diagnosis) scored below cut-off on the SCQ and was excluded from analyses.

ASD participants were recruited from two residential schools: one specialising in Asperger syndrome in South West England and one for children and adolescents with special educational needs in South East England. Only males were recruited because of the preponderance of males in autism (Rutter, 1978) coupled with the male advantage found on visual spatial tasks in the typical population (Maccoby & Jacklin, 1974; McGee, 1979); an exploration of sex differences on coherence measures in autism was beyond the scope of the study. The head teacher gave permission for the study to take place in his/her school, and together with the class teachers, selected suitable children with a known diagnosis of ASD to take part. As with the TD study, letters were sent home to parents and guardians outlining the aims of the study and the nature of their child's participation.

Current IQ data from a full WISC-III or WAIS-III were available or collected by the researcher for 13 participants in the ASD group. Due to time constraints, 18 participants were administered a shortened version based on four subtests (as for the TD sample, Section 4.2). The use of short-forms as an estimate of IQ in ASD populations has recently been validated by Minshew, Turner, and Goldstein (2005). Where available, confirmation of the short-form IQ was made by comparing individual scores with previously collected full-scale IQ data. In all cases it was ensured that that the two IQ scores did not deviate by more than 10 points. Participants with a diagnosis of autism were equivalent in age to those with Asperger syndrome ($t_{(29)} = 0.94, p = .36$), but scored significantly lower on all IQ indices (FIQ 72 vs. 98; VIQ 77 vs. 101; PIQ 74 vs. 95; FIQ-BD 75 vs. 100; all $t_{(29)} > 2.40$, $p < .03$). Within the Asperger syndrome subgroup, four participants had co-morbid attention-deficit/hyperactivity disorder (ADHD) and one had co-morbid attention-deficit disorder (ADD). Three participants had been prescribed medication for the management of their ADHD and were on medication during the experimental sessions. The five participants with ADHD/ADD did not differ in age ($t_{(23)} = 1.00$, $p = .33$) or IQ (all $t_{(23)} < 1.30, p > .20$) from the remaining participants with Asperger syndrome. Examination of their experimental test data also suggested their inclusion in this group did not affect the overall pattern of results. Furthermore, group results for the ASD sample did not alter when these five participants were excluded, and they are therefore included in the analysis reported in this thesis.

Task performance in the ASD sample was compared to that of a control group, comprising 31 males aged from 9 to 20 years. Twenty-six participants from the TD sample were used as age- and ability-matched controls. A further nine participants with moderate learning disability (MLD) were recruited from a special educational needs school in South East England in order to provide an adequate match to the lower functioning ASD participants (four MLD participants were later excluded as they did not provide an adequate IQ/age match). Individuals were excluded from the study if they had fragile-X syndrome or any suggestion of an ASD. As a screening procedure, parents of MLD participants were asked to complete the SCQ. MLD participants were excluded if they scored at or above the 15-point cut-off. Two boys with complex speech and language disorders did exceed the criteria (SCQ scores of 15 and 18). When approximating the scores in the SCQ to the ADI-R algorithm (by converting yes responses to an ADI-R score of 2), neither individual scored beyond cut-off on all three domains (reciprocal social interaction, language/communication, restricted, repetitive and stereotyped behaviours) required for a diagnosis of an ASD (Le Couteur et al., 1989). The two individuals therefore remained as control participants. Furthermore, examination of experimental test data suggested inclusion of these participants did not affect the pattern of results.

Control participants were administered the short-form of either the WISC-III or WAIS-III as described above. An attempt was made to individually match control participants with ASD participants on age and ability. Twenty-two participant pairs were matched in age within one year, and 26 pairs were matched within 10 standard score FIQ points. Exceptions occurred for five low-functioning (FIQ < 70), two very high-functioning (FIQ > 130), and two older participants (age > 18 years), where it was difficult to find equivalent matches in both age and FIQ. Age and IQ data for each ASD participant and their respective matched control are presented in Appendix B.

Group characteristics for ASD and matched control participants are presented in Table 4.3. Age and IQ data are also presented for groups split by intellectual ability. Participants with FIQ of at least 75 comprised the high IQ group (22 Asperger syndrome, 3 autism); participants with FIQ below 75 comprised the low IQ group (3 Asperger syndrome, 3 autism). Statistical comparisons confirmed the ASD and control groups did not differ significantly in age ($t_{(60)} = 0.01, p = .99$) or IQ (all $t_{(60)} < 0.23, p > .82$). This was also confirmed for high IQ (all $t_{(48)} < 0.32, p > .75$) and low IQ subgroups (Mann-Whitney U, all $z < 1.38, p > .16$).

A moderate, although non-significant, negative correlation was found between age and FIQ in the ASD ($r = -.20, p = .28$) and control groups ($r = -.30, p = .10$). Mental age (MA) scores were therefore constructed (based on individual FIQ-BD scores) in order to counteract any age/IQ confounds in the data. No significant group differences were found on MA (all $t_{(60)} = 0.22, p = .83$, high IQ $t_{(48)} = 0.55, p = .58$; low IQ $z = .32, p = .75$).

Table 4.3 ASD and control group characteristics: Mean (*SD*).

	Group	N	Age (years)	MA (years)	FIQ	VIQ	PIQ	FIQ-BD
All	ASD	31	14.8 (2.4)	14.1 (3.8)	93.5 (23.4)	96.7 (22.9)	91.2 (21.2)	95.5 (22.5)
	Control	31	14.8 (2.4)	13.9 (3.2)	94.1 (21.0)	97.1 (21.8)	92.4 (17.8)	94.7 (20.8)
High IQ	ASD	25	14.7 (2.5)	15.2 (3.2)	102.2 (16.5)	104.2 (18.3)	98.9 (15.4)	103.7 (15.9)
	Control	25	14.5 (2.4)	14.7 (2.5)	101.8 (14.3)	104.4 (16.2)	98.4 (13.1)	102.4 (13.8)
Low IQ	ASD	6	15.3 (1.6)	9.3 (1.5)	57.5 (6.8)	65.1 (7.6)	59.3 (6.1)	61.1 (8.0)
	Control	6	16.1 (2.2)	10.3 (3.1)	61.9 (11.1)	66.7 (14.6)	67.4 (12.4)	62.8 (13.3)

IQ profiles in ASD traditionally show peaks and troughs in performance across subtests (e.g., Joseph, Tager-Flusberg, & Lord, 2002). In the present study, however, no significant differences were found between the ASD and matched controls on the individual subtests from the WISC-III and WAIS-III (all $t_{(60)} < 0.59$, $p > .55$, see Table 4.4). Of note is the absence of superior performance on the Block Design subtest in the ASD group compared to the control group. Indeed, only 11 ASD participants (8 Asperger syndrome, 3 autism) were found to show peak Block Design relative to their own Performance Scale score. This was defined by a Block Design subtest score that was at least one point higher than their mean scaled score based on the Performance scale subtests, after excluding Block Design (either full or prorated score). ASD participants with peak Block Design were somewhat younger than ASD participants who did not show peak Block Design (13.9 years vs. 15.3 years), but not to a statistical significant degree ($z = 1.86$, $p = .06$). No differences in IQ were observed between peak and non-peak Block Design ASD participants (all $z < 0.91$, $p > .35$).

Table 4.4 ASD and control group characteristics on Wechsler subtests: Mean scaled score (*SD*)

	Group	N	Information	Vocabulary	Picture Completion	Block Design
All	ASD	31	9.2 (4.8)	9.5 (3.7)	9.0 (3.8)	8.5 (4.1)
	Control	31	9.6 (4.9)	9.4 (3.6)	8.5 (3.2)	8.9 (3.8)
High IQ	ASD	25	10.7 (4.1)	10.8 (3.0)	10.3 (2.9)	10.0 (2.9)
	Control	25	11.1 (4.0)	10.6 (2.6)	9.5 (2.5)	10.0 (3.1)
Low IQ	ASD	6	3.2 (1.8)	4.5 (1.8)	3.7 (1.9)	2.3 (1.8)
	Control	6	3.3 (2.7)	4.3 (2.9)	4.5 (2.3)	4.3 (2.6)

4.3 General Procedure

Ethical permission for the study on local-global processing and cognitive style in ASD and TD was obtained from the Institute of Psychiatry/South London and Maudsley NHS Trust Ethical Committee (Study No. 034/99). Written informed consent was obtained from a parent or guardian for every school-aged participant, while all adult participants gave their own written informed consent to take part.

4.3.1 Testing procedure

All school-aged participants were tested individually by the researcher in a quiet room at school, while adult participants were tested individually at the Social, Genetic and Developmental Psychiatry (SGDP) Centre in South London, or at the Psychology Department, University of Auckland. Participants were given a brief verbal explanation of the study before the first session began. They were informed that the study was looking at individual differences in problem-solving styles, without emphasising the nature of local-global processing. Participants were also told that they were free to leave at any time, without having to provide an explanation. Only one ASD participant (with Asperger syndrome) refused to return for the final test session, which resulted in missing data for all or part of 10 tasks.

All participants were assessed on the coherence task battery consisting of 20 novel and adapted measures of local-global processing (see Section 4.4). The testing sessions also included the short-form of the WISC-III or WAIS-III, and self-report questionnaires on state and trait anxiety, depression, and real-life detail-focus. Results of the questionnaire measures are beyond the scope of the present thesis, and will be reported elsewhere. Three sessions of approximately 60 minutes each were required to complete the entire battery for school-aged participants, and two sessions of 90 minutes for adult participants. For the majority of TD participants, the sessions took place within a two-week period (n = 143) or a month (n = 52). In exceptional cases, the sessions were spaced within a two-month period (n = 9). Of the 40 older TD participants not seen in a school setting, 35 chose to complete the sessions in one day for convenience. In these situations a 30 to 60 minute break was taken between sessions to prevent fatigue. For the majority of the ASD and control participants, the three testing sessions were spaced within a two-week period (n = 41) or a month (n = 16). Due to the limited access to the specialist school in South West England, the sessions there were spaced within a three-month period for four ASD participants. A seven-month gap transpired for one ASD participant who unfortunately broke his arm after the second session.

Eleven experimental tasks were presented using the software SuperLab Pro (Version 2.0) through a Toshiba laptop computer (Satellite Pro 4200). Stimuli were presented on a 15-inch desktop touch screen (Elo Entuitive 1525L LCD), which also served as the response input module for nine experimental tasks. The response input for the two remaining computer-administered tasks was a two-button serial mouse that enabled one millisecond resolution in timing. It was observed that the MLD participants with motor co-ordination problems had difficulty manipulating the mouse buttons for two computer-administered tasks. To remove this potential confound with task performance, these participants verbalised their answer and the experimenter pushed the corresponding buttons. All onscreen instructions were read aloud by the researcher in order not to disadvantage those with low reading ability.

The order of administration for all measures in the task battery, across three sessions, is presented in Appendix C. The order was maintained when the testing was split into two sessions. Tasks alternated between pencil-and-paper and computer-administered, as well as between visual and verbal modalities, in order to provide variety. Some exceptions occurred to the set order when extra time became available or when participants were slower on certain tasks than anticipated. The order was disrupted for three TD participants after the laptop computer malfunctioned, and pen-and-paper tasks could only be administered in one session. Positive comments were made throughout the sessions to encourage participants, but no feedback was given about the correctness of responses during the test phase of a task.

4.4 The coherence task battery

Following a literature review of existing measures of local-global processing in diverse fields of psychology (e.g., developmental, individual differences, neuropsychology), tasks for the current study were selected to span modalities (visual, verbal/auditory) and levels of processing (perceptual, strategic, semantic). New tasks were designed to fill gaps identified in existing research. Extensive piloting (90 participants, aged 8 to 25 years) resulted in 20 suitable tasks in the finalised coherence battery, including nine newly developed tasks, seven from previous studies of central coherence in ASD (Happé, 1997; Happé & Booth, 2006; Heaton, 2003, 2006; Heaton et al., 1999; Scheuffgen, 1998; Shah & Frith, 1983), and four adapted from research in local-global processing in TD (Akshoomoff & Stiles, 1995a; M. J. Cohen, 1997; Dukette & Stiles, 1996; Fodor & Bever, 1965). This thesis presents data from 14 tasks (further tasks, not included due to space constraints, will be reported elsewhere). See Table 4.5 for titles of the tasks, and Chapters 5 through to 8 for details. The effect size (Cohen's **d**; Howell, 2002) for ASD and control group comparisons on each coherence measure is presented in Appendix Q (ranked in order of magnitude). Inter-rater reliability was established for all subjective tasks and this information is presented in the respective Appendix.

Table 4.5 Tasks in coherence battery, divided by modality and level of processing

	VISUO-SPATIAL TASKS	AUDITORY-VERBAL TASKS
HIGHER-LEVEL TASKS	Drawing Task: (1) drawing style (2) effect of meaning on recall Picture Memory: (1) global recall (2) surface form recognition Fragmented Pictures	Sentence Completion Homograph Reading Story Memory: (1) verbatim recall (2) sentence recognition
LOWER-LEVEL TASKS	Un/segmented Block Design Embedded Figures Test Categorisation of Im/possible Figures Navon Similarity Judgment	Phoneme Segmentation Chord Segmentation Pitch Identification Chord Sequence

4.5 Language level

As the coherence battery included several language tasks, information was gathered from each participant on his/her experience with the English language. The frequency of participants who had English as their only language, English as their first language (but who spoke another language fluently), or English as a second language (ESL) is presented in Appendix D. Removing the twelve TD participants with ESL from analyses did not alter the pattern of significant age group and gender differences on the IQ indices. Furthermore, removing the one control participant with ESL (12 years, high IQ group) did not alter the age and IQ matching between clinical and control groups. As language level may impact performance on high-level verbal measures of central coherence, group comparisons in the TD sample were examined with and without participants with ESL.

4.6 Music experience

As the coherence task battery included three tasks that depend on musical abilities (Pitch Identification, Chord Segmentation, Chord Sequence), information was gathered from each participant on formal music training. The number of years of music training received by each participant was then used as a covariate on the music measures. Detailed information on the absolute number of years of music training by age group, and age groups split by years of music training (less than one year versus one year or more), can be found in Appendix E (along with statistical analyses). In sum, music experience was associated with higher IQ and age in the TD sample. Although no main effects of gender were observed, females in the 14-16 year group had received significantly more years of music tuition than males in this age group. Control participants had received significantly more years of music training than participants in the ASD group, particularly when comparing participants with high IQ. Five ASD participants had received formal music training for more than one year, compared to 11 control participants. These two subgroups were matched in age, however the control group was found to have higher FIQ and VIQ than the ASD musically trained group.

4.7 Statistical analyses

The Statistical Package for Social Science, version 12.0 for Windows (SPSS, Inc., Chicago, Illinois) was used to analyse the data presented in this thesis. Initial exploratory analyses were conducted to examine the normality of the data for each coherence task. Although Tabachnick and Fidell (1996) state that it is acceptable to use parametric statistics when the sample size is greater than 200, a conservative approach was taken and non-parametric procedures were followed when violations of normality occurred.

Analyses were conducted with and without outliers (data points exceeding 3 *SD* above/below group means) and if no difference occurred, outliers remained. When data were considered to violate assumptions of normality, non-parametric statistical tests were used (Kruskal-Wallis Test χ^2, Mann-Whitney U, Wilcoxon's Signed Ranks Test, Fisher's Exact Test, Spearman's r_s as appropriate). In all Mann-Whitney U test results, z scores are reported for clarity and p values corrected for ties, where appropriate.

As recommended by Streiner (1996), error bars on graphs show 95% confidence intervals (plus and minus 1.96 standard errors) in order to indicate the range of values within which the "true" value lies. Error bars were not included when they would obscure the clarity of the graph. Boxplots were used to show the distribution of data by quartiles and the median. In each case, the ends of the whiskers represented the largest and smallest values that were not outliers (i.e., within 1.5 times the interquartile range).

One-tailed tests with $p \leq .05$ were used whenever a priori predictions were made (e.g., ASD vs. controls on measures of central coherence). Developmental effects were predicted on several coherence measures, with the expectation of improvements in performance with age and IQ. One-tailed tests were therefore considered appropriate to test for predicted directional effects. However, as statistical analyses across the four age groups in the TD sample required adjustment for multiple comparisons (Howell, 2002), two-tailed tests were used throughout for consistency (with $p \leq .05$).

Since a large number of potential variables were generated by the 14 tasks given to 6 groups, data reduction was attempted when possible. In order to examine performance across tasks, key variables were selected for each task. The key variables (both categorical and continuous) were decided upon through various factors: ability to discriminate ASD and control group performance, show substantial individual variation in all groups, and capture an adequate number of participants showing weak coherence in order to make meaningful comparisons.

Multiple regressions were used on each index of coherence to obtain the proportion of variance not explained by age, IQ, and gender, which may reflect individual differences in cognitive style. The adjusted R^2 value was used to indicate the proportion of the variance in each coherence index that is accounted for by the model. Multiple regressions were performed on the larger TD sample only, as this procedure is considered unsuitable for small samples, particularly when the distribution of scores is very skewed (e.g., $N > 50 + 8m$, where m = number of independent variables; Tabachnick & Fidell, 1996, p. 132).

Chapters 5 through to 8 present group data (group, age, gender, and IQ effects) on coherence tasks in each of the four quadrants (Table 4.5). Chapter 9 presents an analysis of performance across tasks, within and then across quadrants.

Chapter 5. Low-level visuo-spatial coherence tasks

5.1 Introduction

Chapter 5 presents experimental findings for four visuo-spatial tasks designed to tap central coherence at a low level of processing. Tasks were selected to balance the demands for local/global processing; thus a drive for local processing predicts good segmentation ability (EFT, Un/segmented Block Design), while a predominant global processing style predicts good visual integration (Impossible Figures). Evidence of detail focus in ASD may therefore be demonstrated by not only superior performance on visuo-spatial measures that demand local processing but also poor performance on tasks that demand global processing. The preference for global over local-detail is also assessed within the same task (Navon Similarity Judgment Task).

An introduction is provided for each coherence measure, followed by task design, administration, developmental results from the TD sample, and comparison between ASD and age- and IQ-matched controls. Coherence tasks were either based on findings in the ASD or TD literature, and as such, the introduction precedes with the more relevant source. This format follows for each coherence measure presented in the subsequent three chapters, which represent each processing-level/modality quadrant.

5.2 Un/segmented Block Design Task

The Block Design test was originally developed by Kohs in (1923) as a measure of general intelligence. Since its inception it has become a well-established subtest of the Wechsler scales with the highest loading on *g* in the Performance scale (Wechsler, 1992). In the standard Wechsler version, participants are asked to use red and white coloured blocks to reproduce two-dimensional square, red and white designs. The task is widely viewed as a measure of nonverbal reasoning, perceptual organisation, psychomotor speed, and visual-motor coordination (Sattler, 1992), although two key processes are often described: *analysis*, the process of segmenting a whole into parts, and *synthesis*, the ability to assemble parts to form a whole. A detail-focused processing style is predicted to be advantageous in the ability to analyse a design into its constituent parts. A measure of the relative effect of pre-segmenting the design on construction times was therefore used in the present study as an index of central coherence.

Block Design performance in ASD

Despite the high loading on nonverbal intelligence in the general population, performance on the Block Design subtest is often inconsistent with general ability in individuals with ASD. Indeed peak performance on the Block Design, relative to other subtests, is frequently reported in the ASD literature (e.g., Bartak, Rutter, & Cox, 1975; Happé, 1994b; Lincoln, Allen, & Kilman, 1995; Lockyer & Rutter, 1970; Rumsey, 1992; Venter, Lord, & Schopler, 1992).

To explain this exceptional performance, Frith (1989) proposed that individuals with autism benefit from taking a local, rather than global, approach to solving the Block Designs. In doing so, these individuals were readily able to parse the target pattern into single units, and then reconstruct the pattern, block by block. Shah and Frith (1993) found the relative benefit from presenting designs in a segmented form, compared to the usual whole form, was reduced in autism compared to TD and MA-matched controls. Individuals with autism appeared to immediately see the whole design in terms of its constituent parts and this "enhanced ability to segment a gestalt" (Shah & Frith, 1993, p. 1362) was a proposed consequence of weak central coherence. No group differences emerged in other conditions, such as the effect of obliqueness or rotation of the design, which might have indicated that Block Design superiority was due mainly to general visuo-spatial superiority.

Recent studies have explored the finding of superior performance on the Block Design test in ASD in relation to other factors, such as diagnosis and ability level. Ropar and Mitchell (2001) found that children diagnosed with autism performed higher on the standard Wechsler Block Design subtest than verbally matched (MLD) control participants. However, participants with an Asperger syndrome diagnosis performed at a similar level to age-matched TD controls on this measure. This finding does concur with Shah and Frith (1993) whose participants all had a diagnosis of autism and all performed better than controls on the unsegmented condition, regardless of age and ability.

Pring, Hermelin and Heavey (1995) investigated whether the ability to segment wholes into parts plays a role in artistic talent in individuals with and without autism. Artistically talented participants performed at a superior level on the Block Design test, irrespective of whether they had autism. Even so, individuals with ASD, with and without artistic talent, were still significantly faster in constructing designs than controls matched in MA. The authors concluded that having a diagnosis of autism and/or an aptitude for drawing independently predicted good Block Design performance.

Developmental change in Block Design performance

The standard Wechsler version of the Block Design is a renowned measure of cognitive development, being sensitive to changes during development as well as aging. Normative data obtained for the UK population show that performance on the WISC-III Block Design subtest (Wechsler, 1992) improves until the age of 16. Data on the WAIS-III

Block Design subtest (Wechsler, 1997a) indicate a plateau between the ages of 16 and 34, after which performance starts declining.

Gender differences have frequently been reported on the standard Block Design test, with males outperforming females (Arceneaux, Cheramie, & Smith, 1996; Jensen & Reynolds, 1983; Lynn & Mulhern, 1991). Early reports claimed that the male superiority on visuospatial tasks is not manifested until early adolescence (Maccoby & Jacklin, 1974), although this has not be confirmed for the Block Design subtest. Lynn and Mulhern (Lynn & Mulhern, 1991) for example, found the gender difference was not stronger in older children as compared with younger children (split at 12 years) in a large sample of Scottish children (687 males, 709 females). Jensen and Reynolds (1983) also confirmed that the pattern of gender difference was the same in both younger and older groups in an American sample, again split at 12 years (944 males, 924 females).

Several researchers have attempted to identify factors that may contribute to developmental changes in Block Design performance. Much research has looked into strategy use, with the suggestion that participants typically use one of two distinct strategies in constructing designs (Akshoomoff & Stiles, 1996; Jones & Torgesen, 1981; Schorr, Bower, & Kiernan, 1982). In the *analytic* strategy, the target design is mentally segmented into its constituent parts that correspond to the block faces, and then the blocks are placed, one by one, to match these parts. Alternatively, in the *global* or *synthetic* strategy, the target design is viewed more as a whole and the blocks are manipulated until they match the pattern, typically by a process of trial and error. It has been found that taking an analytic strategy is more efficient than the global strategy in solving the designs (Schorr et al., 1982). Furthermore, an analytic strategy is akin to the approach taken by individuals with autism as described by Shah and Frith (1993).

Akshoomoff and Stiles (1996) studied how children (4.5 to 9 years) constructed designs that differed in perceptual cohesiveness (i.e., the strength of the gestalt pattern). They found that performance and error types differed significantly as a function of age and pattern type. For example, errors that broke the square configuration (2 by 2, or 3 by 3) were more common with globally cohesive patterns and were produced more often by younger children. The authors compared performance when a square grid, which corresponded to the single block edges, was superimposed over the design. In this condition, the children who had difficulty parsing the cohesive patterns were able to modify their strategy and made fewer configural errors. Akshoomoff and Stiles concluded that with development, children improve on Block Design tasks due to employing a "more efficient analytic strategy" (p. 401).

Rozencwajg and Corroyer (2002) also studied the development of strategy use in individuals aged 12 years, 17 years, and 20 to 35 years, in a computerised task derived from the standard Block Design test. The target designs were made appropriately complex and consisted of an obvious gestalt that was difficult to break down into components. Although their participants were older than those in the Akshoomoff and Stiles (1996) study, Rozencwajg and Corroyer found their 12-year-old participants would more often use a global strategy than the older participants in their sample. The use of an analytic strategy was more frequently used in older participants than younger.

Kramer et al. (1999) studied the significance of configural errors made by children (6 to 14 years) on the Block Design subtest. The frequency of configural errors was inversely related to overall performance on the subtest, as well as directly related to a local response bias on a quite separate similarity judgment task (Kimchi & Palmer, 1982). When solving Block Designs, participants who are less influenced by the global properties of a visual stimulus, may concentrate on only a part of the pattern and as a consequence, violate the overall configuration (e.g., by using an incorrect number of blocks). This finding has important implications for the weak central coherence theory: if an individual with weak central coherence does make errors on the Block Design test, these errors should be of a type that preserve the details within a design, but break the configuration.

The Un/segmented Block Design methodology originally devised by Shah and Frith (1993) was used in the current study. In preference to assessing absolute ability, performance is contrasted between two conditions: (1) when the target designs are presented in their typical, unsegmented form, and (2) when the target designs are presented in segmented form. The degree of advantage from the segmentation is taken as a measure of central coherence: the larger the degree of benefit indicating stronger coherence.

Predictions

The ASD literature suggests that proficient Block Design performance is a marker of weak coherence due to an enhanced ability to break down the design into its constituent parts (Shah & Frith, 1993). It was therefore predicted that the ASD group would show a smaller degree of benefit from segmentation relative to age and IQ-matched control participants. Construction errors that break the configuration of the target design were recorded for participants in the ASD and control groups. As configuration-breaking errors have been associated with a local perceptual bias (Kramer et al., 1999), it was expected that they would occur more often in the ASD group compared to the control group.

Performance on the standard Block Design test is known to increase with age and ability, and indeed Shah and Frith (1993) found, when contrasting segmented and unsegmented conditions, participants of low ability showed greater benefit from segmentation than those of high ability; likewise, younger children gained more benefit from segmentation than older children. Developmental effects are therefore predicted in the present study with greater benefit from segmentation in younger individuals and those of low ability. An effect of gender is also predicted, with male participants expected to show an advantage in performance overall and less effect of segmentation.

5.2.2 Method

Materials

Stimulus materials included the set of nine yellow and black blocks (2.5 x 2.5 cm) from the British Ability Scales (Elliot, Murray, & Pearson, 1996). Each block has one yellow side, one black side, two sides that are divided horizontally into yellow and black, and two sides that are divided diagonally into yellow and black.

The target patterns consisted of a set of six unsegmented and six segmented yellow and black designs, as shown in Figure 5.1. In each set, four designs required four blocks to construct each pattern (2 by 2 squares), and two designs required nine blocks (3 by 3 squares).

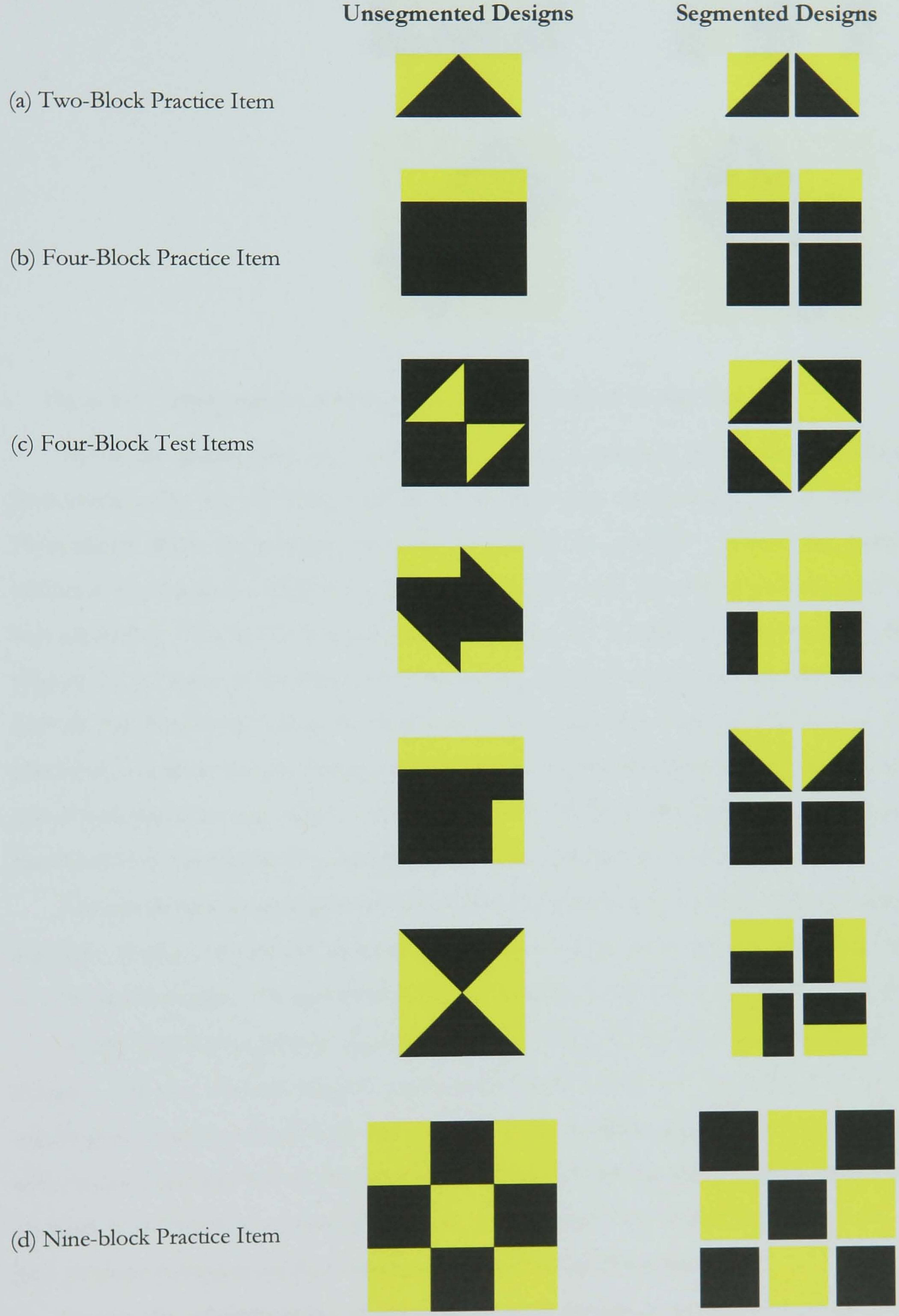
Designs were selected from a set of 31 that had been extensively piloted by the author on mainstream children (6 to 16 years). Data were collected on accuracy and time to construct each design in its unsegmented form, and this information was used to equate the two sets of six designs. Pilot data for the selected items are presented in Appendix F.

The designs appeared centrally on a 15-inch computer touch screen, using SuperLab Pro software controlled by a laptop computer. For unsegmented designs, the dimensions of four-block patterns were 5.5 by 5.5 cm and nine-block designs were 7.6 by 7.6 cm. For segmented designs, the four-block patterns measured 6 by 6 cm and nine-block designs, 9 by 9 cm.

Procedure

Participants were first familiarised with the two-coloured blocks and shown how each block was identical with different coloured sides (black, yellow, half black and yellow horizontally and vertically). The researcher then arranged four blocks to match the four-block practice item (Figure 5.1, b) and explained to the participant how the top of the blocks looked the same as the picture. A rotation error was also demonstrated (i.e., the blocks rotated more than 30 degrees from the horizontal) and the participant was informed this answer would be incorrect.

The blocks were scrambled and the participant was asked to assemble the blocks to match the four-block practice item. The researcher repeated the demonstration if the participant placed the blocks incorrectly, and the participant had the opportunity for a second attempt. If the participant experienced further failure, the two-block practice item (Figure 5.1, a) was administered. This was first demonstrated by the researcher before the participant had the opportunity to construct it.



(e) Nine-block Test Items

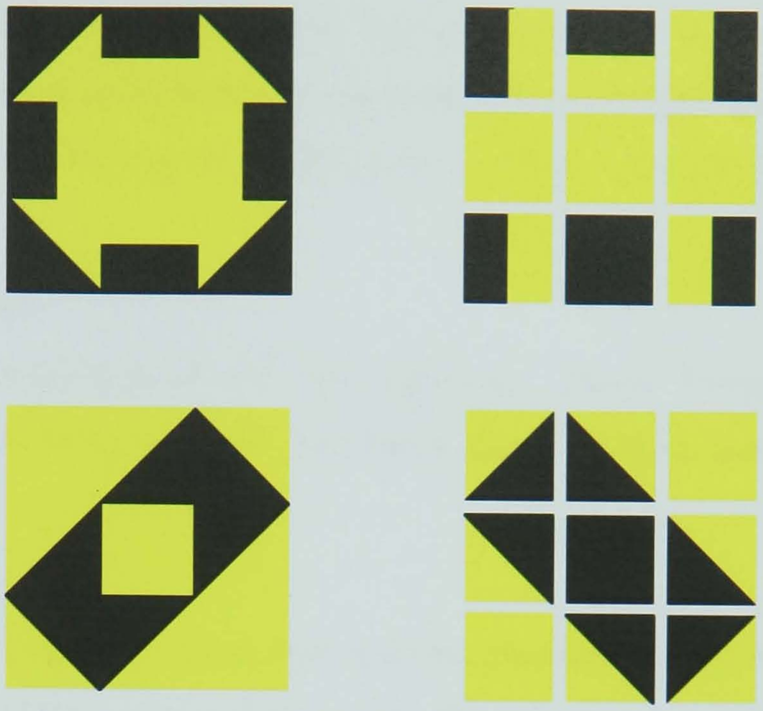


Figure 5.1 Target patterns used in the Un/segmented Block Design Task.

When the participant successfully constructed a practice item (either two-block or four-block), the set of four four-block designs was administered in a fixed order. Participants were encouraged to work as quickly as possible. Once the participant indicated he/she had completed a design, the blocks were scrambled and the next design was presented. The nine-block practice item (Figure 5.1, d) and the two nine-block designs (Figure 5.1, e) were administered to participants who had successfully constructed at least one of the four-block designs. Participants were told that they needed to use all nine blocks to construct the remaining designs and were given the opportunity to construct the nine-block practice item without demonstration from the researcher. Any errors made on the nine-block practice item were highlighted and corrected by the researcher.

The researcher recorded the accuracy of each construction and the time to complete in seconds. A time limit of two minutes was applied to four-block designs and three minutes to nine-block designs. Completions that exceeded these times were recorded as errors.

In the first testing session, participants were presented with the set of six unsegmented designs. In the second session participants were asked to construct the set of six segmented designs, preceded by the corresponding practice items. The unsegmented set was always presented before the segmented set (as in Shah & Frith, 1993) in order to avoid prompting participants to take an analytic strategy; that is, to mentally segment the design into separate components that correspond to individual block faces.

During the administration of the unsegmented designs to participants in the ASD and control groups, the researcher kept a tally of each instance in which the overall configuration of the design was broken. A break in configuration was defined as any instance in which the participant placed a block outside the 2 by 2 or 3 by 3 square matrix

(Kramer et al., 1999). Configural breaks that occurred during construction as well as the final product were recorded. In order to increase the opportunity for configuration breaks to be shown, these were also tallied during the administration of the standard Wechsler Block Design subtest.

5.2.3 Typical Development Results

All TD participants were administered the Un/segmented Block Design Task (N = 204), with all participants receiving both the four-block and nine-block patterns to construct.

Accuracy

Table 5.1 presents the mean number of segmented and unsegmented designs correctly constructed within the time allocated for each age group. As the data were heavily skewed towards correct responses (accuracy ranged from 73% to 98% for unsegmented designs, and 99% to 100% for segmented designs), non-parametric analyses were conducted. A significant effect of age group emerged for unsegmented designs (Kruskal Wallis test, $\chi^2 = 68.0$, $p < .0005$), with accuracy increasing up until the 14-16 year group. No age group effects were found for segmented designs ($\chi^2 = 5.52$, $p = .14$), possibly due to ceiling effects in this condition.

Table 5.1 Number of correct constructions for unsegmented and segmented block designs (maximum = 6) for each TD age group: Mean (*SD*)

Age group (years)	N	Unsegmented designs[a]	Segmented designs
8-10	54	4.35 (1.51)	5.91 (0.35)
11-13	43	5.21 (1.23)	6.00 (0.00)
14-16	51	5.84 (0.50)	5.98 (0.14)
17-25	56	5.88 (0.51)	5.98 (0.13)

[a]8-10 < 11-13 < 14-16, 17-25; $p < .003$

Table 5.2 presents the mean number of correctly constructed unsegmented and segmented designs for male and female participants, overall and in each age group. The predicted male advantage in performance was found with males constructing more unsegmented designs correctly than females ($z = 1.83$, $p = .03$, one-tailed). This gender effect was not evident within each age group however, with only males in the 8-10 year group showing a tendency to outperform females ($z = 1.44$, $p = .07$). No gender differences were found between age groups in constructing segmented designs, which may, again, be a result of ceiling effects in this condition.

Table 5.2 Number of correct constructions for unsegmented and segmented block designs (maximum = 6) for male and female TD participants within each age group and overall: Mean (*SD*)

Age group (years)		N	Unsegmented designs	Segmented designs
8-10	*male*	19	4.68 (1.57)	5.84 (0.50)
	female	35	4.17 (1.46)	5.94 (0.24)
11-13	*male*	21	5.24 (1.26)	6.00 (0.00)
	female	22	5.18 (1.22)	6.00 (0.00)
14-16	*male*	30	5.77 (0.63)	5.97 (0.18)
	female	21	5.95 (0.22)	6.00 (0.00)
17-25	*male*	27	5.89 (0.42)	5.96 (0.19)
	female	29	5.86 (0.58)	6.00 (0.00)
All	*male*	97	5.47[a] (1.08)	5.95 (0.27)
	female	107	5.19 (1.29)	5.98 (0.14)

[a]male > female, p = .03, one-tailed.

Construction time

Mean construction times for unsegmented and segmented designs, irrespective of correctness, are presented for each age group in Table 5.3. This dependent variable was favoured over mean time for correct constructions as the relatively high error rate in the younger groups meant that correct times did not give an accurate representation of the data. The alternative of assigning maximum time to incorrect constructions resulted in the same pattern of results as actual time, so the latter variable was selected for transparency. Mean construction times were markedly positively skewed; hence a logarithmic transformation of the raw data was performed to normalise the distribution of scores. The transformed data (in logarithmic seconds) were used in all statistical analyses.

Table 5.3 Time per item (in seconds) to construct unsegmented and segmented block designs for each TD age group: Mean (*SD*)

Age group (years)	N	Unsegmented designs[a]	Segmented designs[b]
8-10	54	57.2 (16.7)	23.8 (7.0)
11-13	43	38.8 (17.8)	17.3 (4.7)
14-16	51	27.1 (12.0)	13.8 (2.8)
17-25	56	23.1 (12.8)	12.1 (2.5)

[a]8-10 < 11-13 < 14-16, 17-25; $p < .004$
[b]8-10 < 11-13 < 14-16 < 17-25; $p < .008$

A two-way repeated measures ANOVA was performed to explore the effects of age group and gender on mean construction times in segmented and unsegmented conditions. The main effect of age group was significant ($F_{(3,196)} = 72.4$, $p < .0005$) reflecting quicker construction times with development. Although there was no main effect of gender ($F_{(1,196)} = 0.12$, $p = .73$), the interaction between age group and gender reached significance ($F_{(3,196)} = 2.82$, $p = .04$). When the effects of FIQ-BD were covaried, however, this interaction fell below significance ($F_{(3,195)} = 0.90$, $p = .44$).

The main effect of condition was significant ($F_{(1,196)} = 955.95$, $p < .0005$) indicating quicker construction times for segmented compared to unsegmented designs. No condition by gender interaction was found ($F_{(1,196)} = 0.65$, $p = .42$), but the age group by condition interaction was significant ($F_{(3,196)} = 8.62$, $p < .0005$) and survived partialling out FIQ-BD ($F_{(3,195)} = 10.83$, $p < .0005$). This interaction is depicted in Figure 5.2 and indicates that the benefit from segmentation (as shown by the slope in the graph) is greater for the younger age groups. The condition by age group by gender interaction was moderately significant ($F_{(3,196)} = 2.61$, $p = .053$), although it fell below significance when FIQ-BD was taken into account ($F_{(3,195)} = 1.16$, $p = .33$). The interaction, as shown in Figure 5.2, suggests that the diminishing benefit from segmentation with increasing age occurs earlier for females than males.

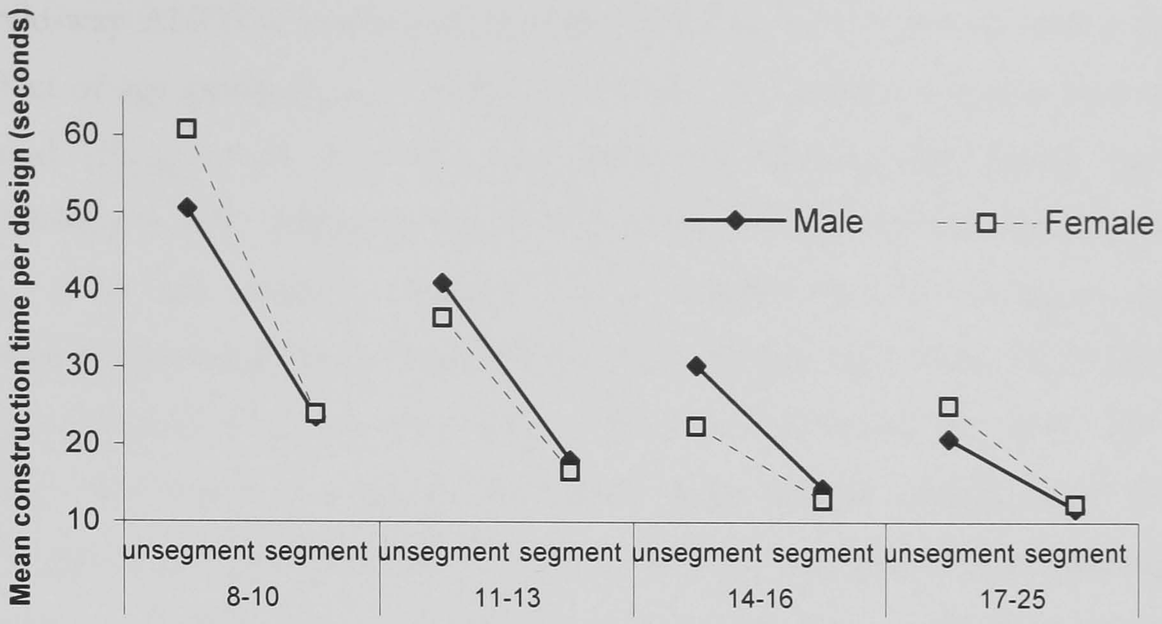


Figure 5.2 Mean construction times (in seconds) for unsegmented and segmented block designs for male and female participants in each TD age group.

Relative Benefit from Segmentation

In order to measures an individual's relative benefit from segmentation, a score was calculated based on the mean difference in construction time (non-transformed) between conditions (unsegmented minus segmented), divided by the mean time for unsegmented designs. The distribution of this relative benefit score, expressed as a percentage, for each age group is presented in Figure 5.3.

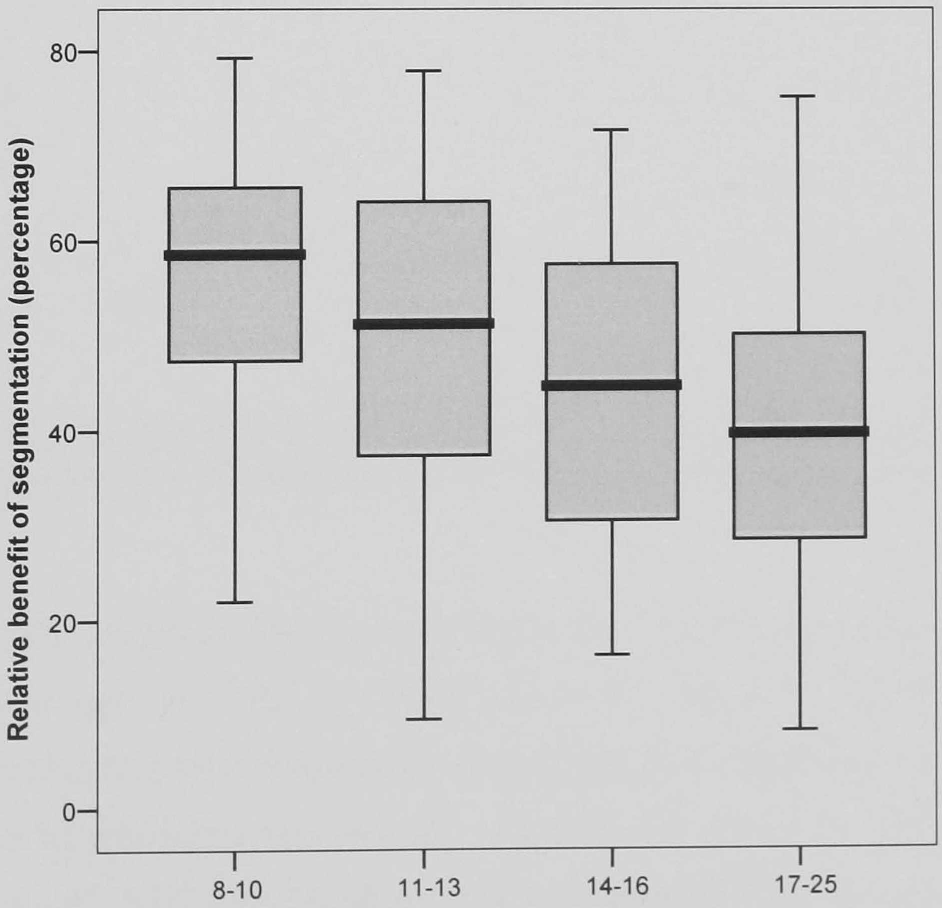


Figure 5.3 Boxplots showing the distribution of the relative benefit from segmentation score for each age group. Each plot shows the median and the distribution of the data by quartiles. The ends of the whiskers represent the largest and smallest values that are not outliers (i.e., within 1.5 times the interquartile range).

A two-way ANOVA performed on the relative benefit scores showed a significant main effect of age group ($F_{(3,196)} = 9.05$, $p < .0005$). No significant main effect of gender was found ($F_{(1,196)} = 0.71$, $p = .40$), nor interaction between age group and gender ($F_{(3,196)} = 2.40$, $p = .07$). Adjusting for the effects of FIQ-BD did not alter the pattern of findings. Post hoc analyses (Tukey's HSD) revealed that the youngest age group demonstrated significantly more benefit from segmentation (M = 56%, SD = 13) than the two oldest groups (14-16 years M = 44%, SD = 15; 17-25 years M = 40%, SD = 17; all $p < .0005$). The 11-13 year group also had a larger relative benefit score (M = 49%, SD = 17) than the 17-25 year group ($p = .03$). Although no interaction between age group and gender was found, Figure 5.4 suggests that the relative benefit from segmentation decreases later for males (between 14-16 and 17-25 years) than for females (between 8-10 and 11-13 years).

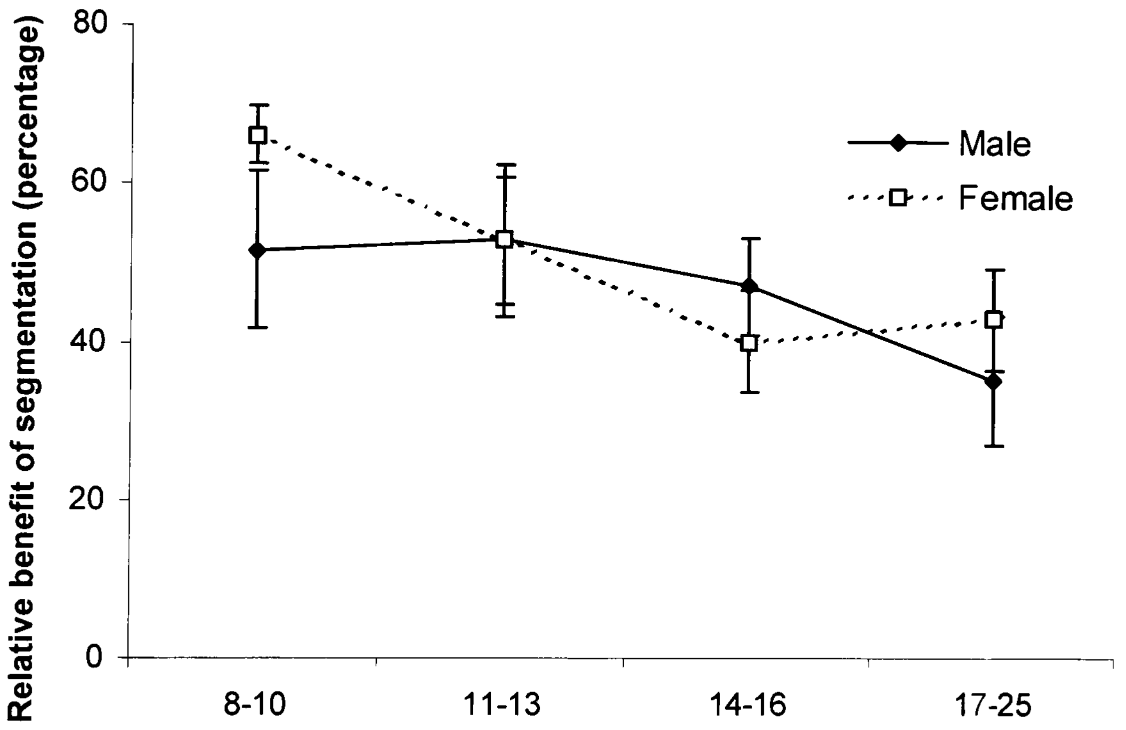


Figure 5.4 Mean relative benefit from segmentation (as a percentage of times) for male and female participants in each TD age group. Error bars show 95% confidence intervals.

Correlations

Across the TD sample, a negative correlation was found between the percent relative benefit score and age (r = -.32, $p < .0005$; males r = -.28, $p = .006$; females, r = -.35, $p < .0005$), indicating that older participants are deriving less benefit from segmentation.

The relative benefit score was strongly related to IQ across the entire sample (FIQ r = -.46; PIQ r = -.47; VIQ r = -.33; all $p < .0005$), suggesting that participants with higher ability were benefiting less from the effect of segmentation. Correlations were significant for male and female participants for FIQ (males r = -.55, females r = -.33, $p < .0005$) and PIQ (males r = -.43, females r = -.20, $p < .0005$), but only in males for VIQ (males r = -.44,

$p < .0005$; females $r = -.18$, $p = .07$; Fisher's z transformation confirmed the correlation coefficients were significantly different between genders $z_{r1-r2} = 2.02$, $p = .04$). The correlation with IQ is unlikely to be an artefact of longer times in general for lower IQ participants, since benefit scores were calculated as a proportion of time for unsegmented designs.

As expected, a strong correlation was evident between the Wechsler Block Design subtest and the relative benefit percentage score ($r = -.58$, males $r = -.70$, females $r = -.47$, all $p < .0005$). Despite this finding, the correlation between ability and relative benefit from segmentation still held when using the IQ estimate that did not include the Block Design subtest (FIQ-BD, all $r = -.33$, $p < .0005$; males $r = -.43$, $p < .0005$; females $r = -.20$, $p = .04$).

5.2.4 ASD and Control Results

One 15-year-old from the control group (with MLD) was not administered the nine-block designs as he did not construct any four-block designs correctly. As his exclusion did not alter the pattern of findings, his average construction time was based on the four-block designs only (segmented and unsegmented). All remaining control and ASD participants were administered the full set of four-block and nine-block designs.

Accuracy

The mean number of segmented and unsegmented designs correctly constructed for the ASD and control groups overall, and split by high and low IQ, is presented in Table 5.4. No significant differences were found between the ASD and control groups for the number of correct constructions in the unsegmented ($z = 0.94$, $p = .18$, one-tailed) and segmented conditions ($z = 1.09$, $p = .14$). Furthermore, no group differences were found for groups divided by IQ, on either unsegmented designs (high IQ $z = 0.85$, $p = .20$; low IQ $z = 0.41$, $p = .35$) or segmented designs (high IQ $z = 1.43$, $p = .08$; low IQ $z = 0.09$, $p = .47$).

Table 5.4 Number of correct constructions for unsegmented and segmented block designs (maximum = 6) for ASD and control groups: Mean (*SD*)

	Group	N	Unsegmented designs	Segmented designs
All	ASD	31	4.81 (1.56)	5.77 (0.56)
	Control	31	5.03 (1.72)	5.84 (0.64)
High IQ	ASD	25	5.24 (1.30)	5.88 (0.44)
	Control	25	5.48 (1.08)	6.00 (0.00)
Low IQ	ASD	6	3.00 (1.26)	5.33 (0.82)
	Control	6	3.17 (2.64)	5.17 (1.33)

Construction time

Mean construction times for unsegmented and segmented designs are shown for each group in Table 5.5. A one-way repeated measures ANOVA found a main effect of condition ($F_{(1,60)}$= 247.4, $p < .0005$) indicating that unsegmented designs took significantly longer to construct than segmented designs. The interaction between condition and group was not significant ($F_{(1,60)}$= 0.14, $p = .71$) suggesting that the effect of condition on construction time was identical in each group. Furthermore, no main effect of group was found ($F_{(1,60)}$= 0.95, $p = .34$) indicating that groups took the same amount of time to construct unsegmented and segmented designs. The same pattern of results was found when FIQ-BD was partialled out of the analyses, and when comparisons were made including only high IQ or low IQ participants.

Table 5.5 Time per item in seconds to construct unsegmented and segmented block designs for ASD and control groups: Mean (*SD*)

	Group	N	Unsegmented designs	Segmented designs
All	ASD	31	41.9 (19.4)	20.3 (9.4)
	Control	31	39.0 (19.0)	17.8 (7.2)
High IQ	ASD	25	38.5 (19.4)	19.1 (9.6)
	Control	25	34.5 (15.0)	16.1 (4.5)
Low IQ	ASD	6	55.9 (12.3)	25.5 (6.7)
	Control	6	58.1 (23.3)	25.0 (11.8)

Relative Benefit from Segmentation

The percentage relative benefit from segmentation (i.e., the difference in construction time between unsegmented and segmented conditions, divided by the time for unsegmented designs) was calculated for each participant. The distribution of these scores for the ASD and control groups is presented in Figure 5.5. No significant group difference was found on this index (ASD M = 47%, SD = 19; control M = 49%, SD = 15; $t_{(60)} = .58$, $p = .29$, one-tailed). This was also the case for high IQ participants (ASD M = 45%, SD = 19; control M = 48%, SD = 16; $t_{(48)} = .50$, $p = .31$), and low IQ participants (ASD M = 52%, SD = 17; control M = 55%, SD = 14; $z = .16$, $p = .44$).

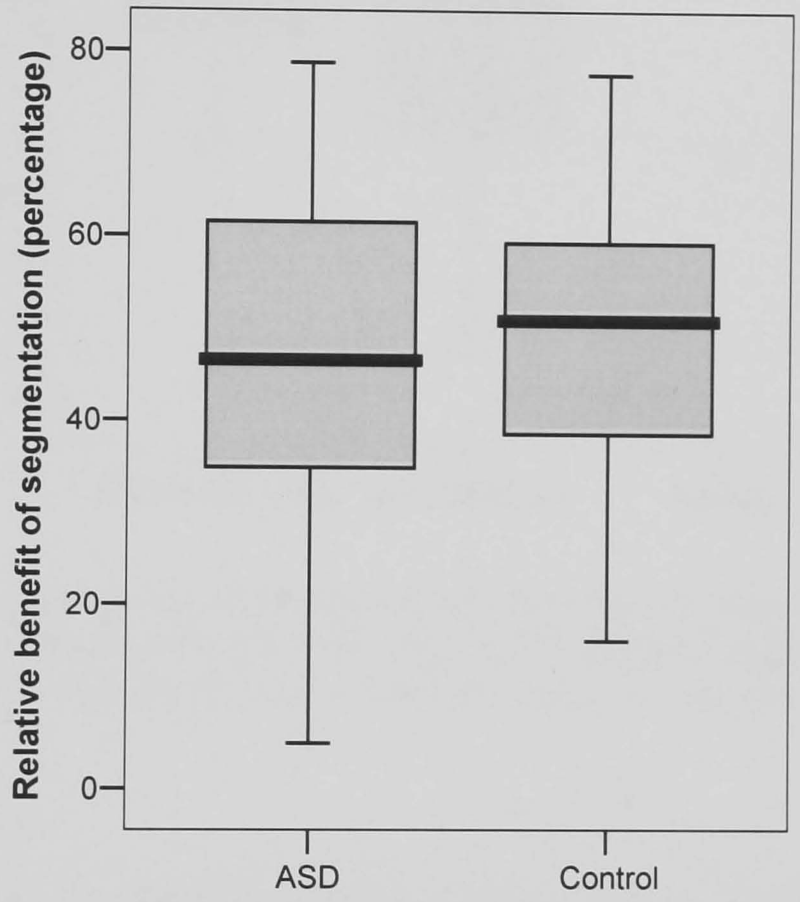


Figure 5.5 Boxplots showing the distribution of the relative benefit from segmentation score for each group. Medians, ranges, and 25 and 75 percentiles are shown.

Subgroup analysis

As discussed in Chapter 4, 11 ASD participants were found to have superior Wechsler Block Design subtest score relative to their own Performance Scale score. Participants in this subgroup were compared to their pairwise-matched controls, as well as to the remaining ASD group who did not show peak performance on the Block Design subtest. The distribution of relative benefit from segmentation for each group is presented in Figure 5.6. The ASD participants with peak Block Design performance derived significantly less benefit from segmentation than their matched controls (ASD M = 41%, SD = 23; control M = 55%, SD = 14; $z = 1.81$, $p = .04$, one-tailed). No significant difference was found between ASD participants with versus without a Block Design peak, perhaps due to large variance in the data from both groups (ASD without peak $M = 50\%$, $SD = 16$; $z = 1.03$, $p = .30$, two-tailed).

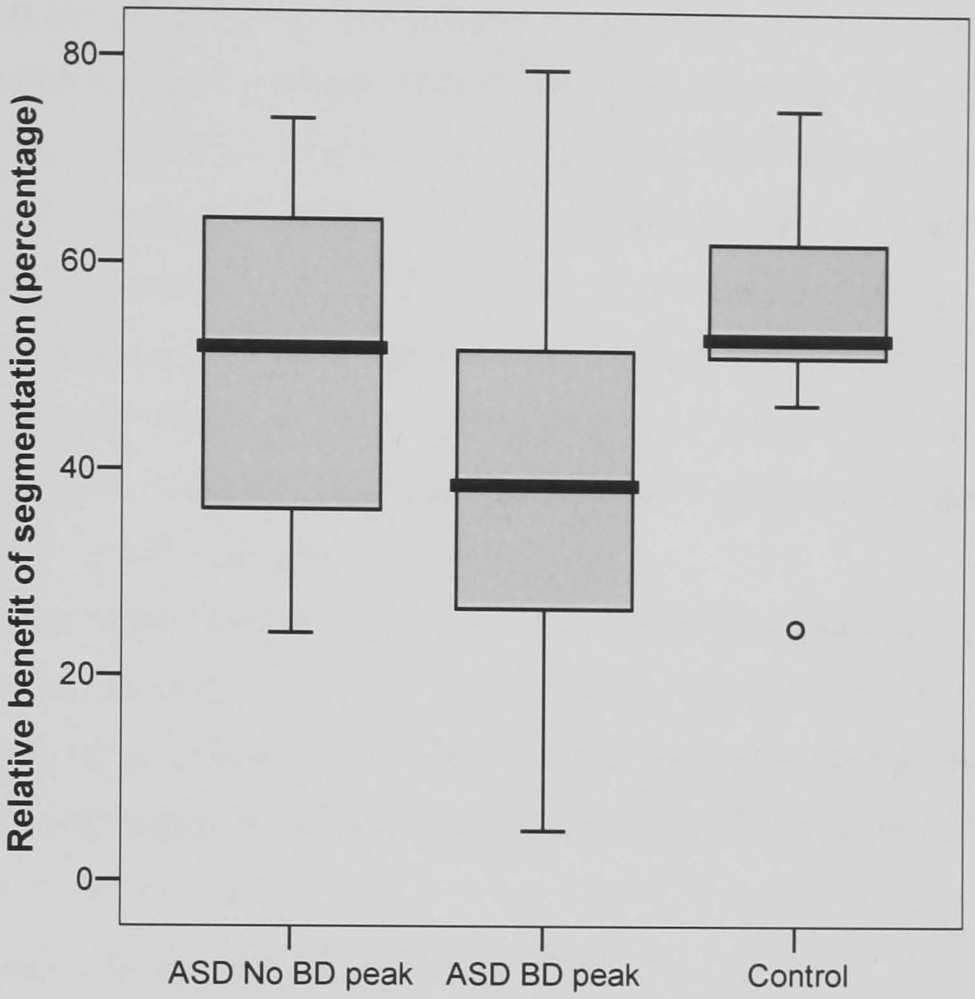


Figure 5.6 Boxplots showing the distribution of the relative benefit from segmentation score for ASD without peak Block Design (N = 20), ASD with peak Block Design (N = 11), and the corresponding matched controls (N = 11). Medians, ranges, and 25 and 75 percentiles are shown.

Configuration-breaking errors

The mean number of configuration-breaking errors observed during construction of the unsegmented designs and the Wechsler Block Design subtest did not differ between ASD ($M = 1.13$, $SD = 1.53$) and control groups ($M = 1.00$, $SD = 1.55$; $z = 0.66$, $p = .33$, one-tailed). Furthermore, no significant differences were detected for groups divided by IQ (high IQ ASD $M = 1.00$, $SD = 1.32$; control $M = 0.72$, $SD = 1.06$; $z = 0.81$, $p = .21$; low IQ ASD $M = 1.67$, $SD = 2.25$; control $M = 2.17$, $SD = 2.64$; $z = 0.42$, $p = .34$). ASD participants with peak Block Design did not differ in the mean number of configuration-breaking errors made ($M = 1.10$, $SD = 1.60$) from their matched controls ($M = 1.00$, $SD = 1.41$; $z = 0.15$, $p = .44$). This subgroup also did not differ from the ASD participants without Block Design peak ($M = 1.15$, $SD = 1.53$; $z = 0.09$, $p = .93$, two-tailed).

Correlations

The relative benefit score was found to have a significant positive correlation with age in the ASD group ($r = .37$ $p = .04$). This correlation was in the opposite direction to the TD sample, indicating that younger participants were deriving less benefit from segmentation. The correlation between the relative benefit score and MA was not significant in the ASD group ($r = .05, p = .77$). No reliable correlation with age or MA was evident in the control group ($r = .19, p = .31$ and $r = -.11, p = .56$, respectively).

A significant negative correlation between the relative benefit score and FIQ was apparent in both groups (ASD $r = -.37, p = .04$; controls $r = -.39, p = .03$). This was also the case for PIQ (ASD $r = -.36, p = .05$, controls $r = -.47, p = .007$) but not significantly for VIQ (ASD $r = -.33, p = .07$, controls $r = -.26, p = .16$).

The only subtest of the Wechsler scales to correlate with the relative benefit score was the Block Design subtest (ASD $r = -.39, p = .03$, controls $r = -.65, p < .0005$). Indeed, the correlation between IQ and relative benefit from segmentation was not significant in either group when the Block Design subtest was not included in the FIQ estimate (ASD $r = -.24$, $p = .19$; control $r = -.25, p = .17$).

5.2.5 Summary of Un/segmented Block Design Task

The predicted age and IQ effects on the Un/segmented Block Design Task performance were confirmed in the TD sample. Younger participants gained greater benefit from the segmentation of the designs than older participants. Furthermore, individuals with lower IQ showed greater advantage from the segmentation than individuals with higher IQ. The relationship between test performance and IQ remained when the Wechsler Block Design subtest was not included in the calculation of FIQ. Gender effects were found in accuracy with males constructing significantly more unsegmented designs correctly than females. Gender effects were not apparent in the relative benefit of segmentation although, graphically, the data suggest the decrease in benefit from segmentation with age occurred earlier for females than males.

ASD participants did not demonstrate the predicted superiority in performance on the Un/segmented Block Design Task, as shown by less benefit from segmentation, compared to their age- and IQ-matched controls. This finding is considered alongside the fact that the majority of the ASD participants did not show personal peak performance on the Wechsler Block Design subtest, commonly found in this clinical population. The subgroup of participants with ASD who did show peak Block Design performance, however, derived significantly less benefit from segmentation than their matched control participants. ASD participants did not show a greater frequency of configuration-breaking errors than control participants, which was considered another potential indicator of detail-focus on this task.

The key dependent variable taken from the Un/segmented Block Design Task for further analysis was the relative benefit from segmentation as a percentage of construction time for each individual, with low scores indicating weaker central coherence. In categorical analyses, participants were categorised as showing weak central coherence if relative benefit score was less than 30%.

5.3 Embedded Figures Test

The Embedded Figures Test (EFT; Witkin et al., 1971) is a commonly used perceptual test of cognitive style and analytic ability. The task involves locating a simple figure within a larger complex form that has been designed to obscure or embed the simple figure. In the children's version of the task (CEFT; Karp & Konstadt, 1963) the complex form is a meaningful picture, while in the EFT the complex form consists of a non-meaningful geometric design. Successful performance is dependent on the ability to resist the tendency to see only the global form or be drawn in by the surrounding context. Lezak (1995) classifies the EFT as a task of visual interference in which the individual is required to "analyse the figure-ground relationship in order to distinguish the figure from the interfering elements" (p. 413). Success on the task is also said to relate to executive control, such as perceptual flexibility and the ability to shift attention between the features of the design. Brosnan et al. (2002), for example, claimed that successful performance on the EFT requires "effective inhibition of the distracting context" (p. 2145).

The EFT was originally developed to study individual differences in cognitive style (field dependence/independence) in typical populations (see Chapter 3) and has become a benchmark measure of central coherence in autism (see Chapter 2). Due to the wealth of research on the EFT it is possible to make predictions of developmental effects and clinical differences expected in the present study. The EFT has shown significant sex differences with males outperforming females in speed and accuracy (Cairns, Malone, Johnston, & Cammock, 1985; Hall, Gregory, Billinger, & Fisher, 1988; Huss & Kayson, 1985). However, some studies have disputed this conclusion with negligible differences found between males and females in older adults (Fogliani-Messina, Fogliani, & di Nuovo, 1983) and in 10 to 12-year-old children (Mahlios & D'Angelo, 1983). Voyer, Voyer, and Bryden (1995) proposed that sex differences in performance on the EFT are not apparent before the age of 14 years. In a meta-analysis carried out on 59 studies that documented sex effects on the EFT, Voyer et al. concluded that sex effects were absent on the CEFT, while medium effects were observed on the standard EFT with males showing superior performance (Cohen's d for effect size = .42).

EFT performance is predicted to increase with age and intellectual ability in children (Huss & Kayson, 1985; Pennings, 1988; Witkin et al., 1971). Pennings (1988) observed that children developed strategies to enable them to solve the EFT between 5 and 7 years of age. Task performance has been found to decline in older adults, typically from 30 years of age (Comalli, 1965; Schwartz & Karp, 1967), although Lee and Pollack (1978) found that females became slower at detecting the embedded shapes in each successive decade from 20 years.

Predictions

Fast and accurate performance on the EFT is predicted to be a key indicator of weak central coherence, over and above differences in age and intellectual ability. Superior performance on the EFT has previously been demonstrated in autism (Jolliffe & Baron-Cohen, 1997; Ropar & Mitchell, 2001; Shah & Frith, 1983; although see Brian & Bryson, 1996; Ozonoff, Pennington, & Rogers, 1991), and it was predicted that the ASD group would outperform the age and IQ-matched controls in the present study. In the TD sample, EFT performance is expected to increase with age and IQ, as surmised from the TD literature. There has also been some suggestion of gender effects, with male superiority emerging in older age groups. As stimuli type differs between the standard EFT (non-meaningful) and CEFT (meaningful), the two tasks were analysed separately to investigate the effect of meaning on disembedding skill. Although, by design, the CEFT is an easier task than the EFT, the meaningfulness of the complex forms may make the task of ignoring the global form more difficult for control participants compared to ASD participants.

5.3.2 Method

Materials

A modified version of the EFT was used, including seven items from the CEFT (all 'house' shape items: 3, 4, 6, 9, 11, 12, 14) and eight items from Form A of the standard EFT (items: 1, 4, 5, 6, 8, 10, 11, 12). Selected items ranged in difficulty in order to produce variability between participants and avoid ceiling effects across the age levels tested. Test items were presented on laminated cards and each simple form was given to the participant on a transparent sheet.

Procedure

Participants were first shown a simple form and told they would soon see a picture that had this shape hidden within it. They were informed that the hidden shape would be the same size and orientation as the simple form and they had to find it as quickly as possible. The simple form remained in view while the test item was presented. This modified procedure was originally adopted by Shah and Frith (1983; and used by Happé et al., 2001) in order to eliminate the memory requirement of the standard EFT where participants are asked to hold the simple form in memory during their search.

Timing began as soon as the test item (i.e., the complex form) was revealed to the participant and stopped when the participant had indicated they had found the simple shape (either by pointing or announcing success). Response times were measured to the nearest second using a hand-held stopwatch. Participants were asked to demonstrate the position of the simple shape by placing the transparent sheet over the complex figure, or by tracing out the shape with their finger. If they were incorrect the simple form was returned to its position next to the complex figure, the participant was encouraged to look again, and timing was resumed. Participants were given 60 seconds to find the target shape within each test item. If they were unable to locate the figure within this time the item was recorded as a failure and they were given a response time of 60 seconds.

Participants were administered items from the CEFT followed by the EFT presented in numerical order. A practice item preceded each set of CEFT and EFT items to ensure the participant had understood the task. Although research has indicated that the items are not presented in order of difficulty (Hardy, Eliot, & Burlingame, 1987), the task was discontinued after five consecutive failures to prevent causing undue stress to participants who might be struggling with the task.

5.3.3 Typical Development Results

Children's Embedded Figures Test

As the CEFT was originally developed for children aged from 5 to 12 years (Witkin et al., 1971) it was considered appropriate to administer the task to TD participants attending school. Individuals in the oldest age group who were still attending school (N = 15, age M = 17.9 years, SD = 0.4, range 17.2 – 18.5 years) were younger than the school leavers (N = 41, age M = 21.5 years, SD = 2.0, range 17.6 – 25.8 years) but were equivalent on all IQ measures.

One 11-year-old male (FIQ 103, VIQ 89, PIQ 118) was unable to locate any embedded figures in the CEFT, while all other participants were able to locate at least three figures. This participant was deemed an outlier in his age group (performance level three standard deviations below the mean) and was removed from further analysis on this measure. Table 5.6 presents the results for the CEFT for each of the four age groups in the TD sample.

Table 5.6 Children's Embedded Figures Test results by TD age group: Mean (*SD*).

Age group (years)	N	Number correct (max = 7)[a]	Time per item (seconds)[b]
8 - 10	54	5.9 (1.0)	21.5 (9.6)
11 - 13	42	6.1 (1.0)	17.3 (9.3)
14 - 16	51	6.5 (0.8)	12.4 (7.5)
17 – 18	15	6.7 (0.5)	9.7 (4.0)

[a]8-10 < 14-16, 17-18, $p < .001$; 11-13 < 17-18, $p = .04$
[b]8-10, 11-13 < 14-16, 17-18, $p < .01$

Accuracy

As the accuracy data did not meet normality assumptions (z-scores for skewness ranged -4.67 to -1.36), nonparametric statistics were used. Significant age effects were found for the mean number of figures located (Kruskal Wallis Test $\chi^2 = 18.14$, $p < .0005$) with older age groups detecting more figures than younger age groups. No differences were found between male and female participants overall ($z = 0.67$, $p = .25$, one-tailed), and within each age group (all $z < 1.09$, $p > .14$).

Response time

Mean response times underwent natural logarithm transformation to obtain normality and were entered into a two-way ANOVA with age group and gender as the between-group factors. A significant main effect of age group was detected ($F_{(3,154)} = 13.47$, $p < .0005$). Post hoc comparisons revealed that the two oldest groups were significantly faster in locating embedded figures than the youngest two groups (Tukey's HSD; all $p < .01$). No main effect of gender was found ($F_{(1,154)} = 0.48$, $p = .49$), while the interaction between age group and gender approached significance ($F_{(3,154)} = 2.49$, $p = .06$). Figure 5.7 suggests that performance on the CEFT improves later for males (between 11-13 years and 14-16 years) than for females (between 8-10 years and 11-13 years). When the ANOVA was rerun with FIQ-BD as a covariate, the main effect of age group was found to be robust ($F_{(3,153)} = 15.93$, $p < .0005$), but the main effect of gender ($F_{(1,153)} = 0.73$, $p = .39$) and interaction between age group and gender ($F_{(3,153)} = 1.54$, $p = .21$) were not significant.

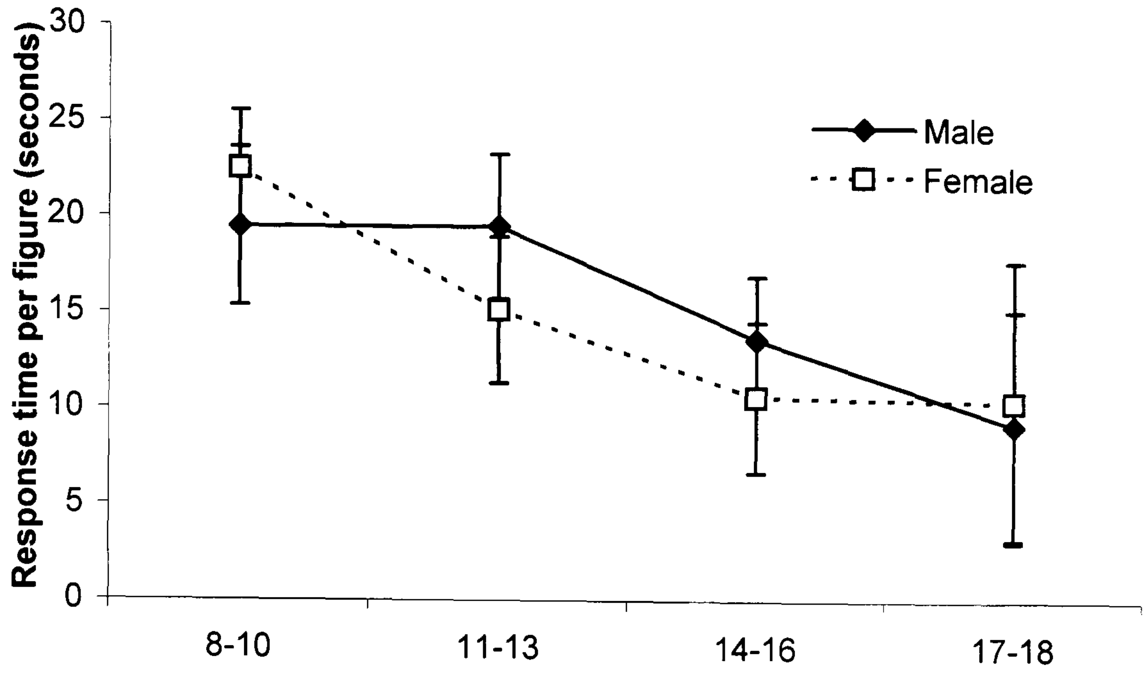


Figure 5.7 Mean response time per figure (in seconds) on the CEFT for male and female participants in each age group. Error bars show 95% confidence intervals.

Correlations

Across the TD participants who were administered the CEFT (n = 162), a strong negative correlation was found between time taken to locate embedded figures and age (r = -.51, males r = -.43, females r = -.58, all p < .0005). Response time was also strongly negatively related to IQ across the entire sample (FIQ r = -.32 p < .0005; FIQ-BD r = -.23, p = .004; PIQ r = -.38 p < .0005; VIQ r = -.19 p < .02). These correlations generally remained significant within each gender (males FIQ r = -.33 p = .003; FIQ-BD r = -.26, p = .02; PIQ r = -.40, p < .0005, VIQ r = -.21, p = .06; females FIQ r = -.31, p = .005; FIQ-BD r = -.19, p = .08; PIQ r = -.37, p = .001; VIQ r = -.16, p = .14), with no difference in the magnitude of the correlation coefficients between males and females (all z_{r1-r2} < 0.47, p > .63).

Standard Embedded Figures Test

All TD participants were administered the standard version of the EFT. Nine participants failed to locate any hidden figures; this group consisted of five females from the 8-10 year group, three males from the 11-13 year group and one male from the 14-16 year group. The latter case was regarded as an outlier for his age group and his data were removed from analyses. The mean number of correct responses and the average time taken to locate an embedded figure are shown in Table 5.7 for each age group.

Table 5.7 Standard Embedded Figures Test results by TD age group: Mean (SD).

Age group (years)	N	Number correct (max = 8)[a]	Time per item (seconds)[b]
8 - 10	54	5.0 (2.3)	35.3 (13.8)
11 - 13	43	6.0 (2.2)	28.6 (12.9)
14 - 16	50	6.6 (1.4)	21.9 (10.0)
17 - 25	56	6.7 (1.5)	21.1 (11.6)

[a] 8-10 < 11-13, 14-16, 17-25, $p < .01$
[b] 8-10, 11-13 < 14-16, 17-25, $p < .03$

Accuracy

The assumptions of normality were violated for the accuracy data (z-scores for skewness ranged -6.26 to -2.23), and nonparametric statistics were used. A significant effect of age group was found for the mean number of EFT figures located ($\chi^2 = 24.83$, $p < .0005$). This group difference was due to the youngest group detecting significantly fewer figures than all older age groups (all $z > 2.33$, $p < .01$). Overall male and female participants did not differ in detection rates ($z = 0.55$, $p = .29$, one-tailed), but gender effects were apparent in the two youngest age groups. In the 8-10 year group, males located significantly more embedded figures ($M = 6.16$, $SD = 1.77$) than females ($M = 4.31$, $SD = 2.35$; $z = 2.93$, $p = .003$). Conversely, in the 11-13 year group, females located more figures ($M = 6.73$, $SD = 1.55$) than males ($M = 5.24$, $SD = 2.47$; $z = 2.42$, $p = .02$). The finding, however, may be an artefact resulting from higher FIQ in males in the 8-10 year group and, since IQ could not be covaried in these non-parametric analyses, this was explored further in the response time analyses.

Response time

A two-way ANOVA was conducted on the time taken to locate each figure (log transformed to obtain normality) with age group and gender as the between group factors. A main effect of age group was detected ($F_{(3,196)} = 12.22$, $p < .0005$) which was accounted for by the two younger groups taking significantly longer to detect the embedded figures than the two older groups. Although no main effect of gender was detected ($F_{(1,196)} < 0.001$, $p = .99$), a significant interaction between age group and gender was found ($F_{(3,196)} = 8.03$, $p < .0005$) and is depicted in Figure 5.8. The pattern of findings from the ANOVA remained after controlling for the confounding effects of FIQ-BD.

The significant age group by gender interaction may reflect a different trajectory of development according to gender, with improvement on the EFT occurring earlier for females than males. To explore this further, one-way ANOVA's were conducted for each gender group. Significant effects of age group on mean response time were found for both males ($F_{(3,92)} = 7.30$, $p < .0005$) and females ($F_{(3,103)} = 16.91$, $p < .0005$). Post hoc comparisons showed that the significant change for females occurred between 8-10 years and 11-13 years ($p < .0005$), while no improvement was apparent after 11-13 years. For males however, an increase in performance was evident between 11-13 years and 14-16 years ($p = .03$).

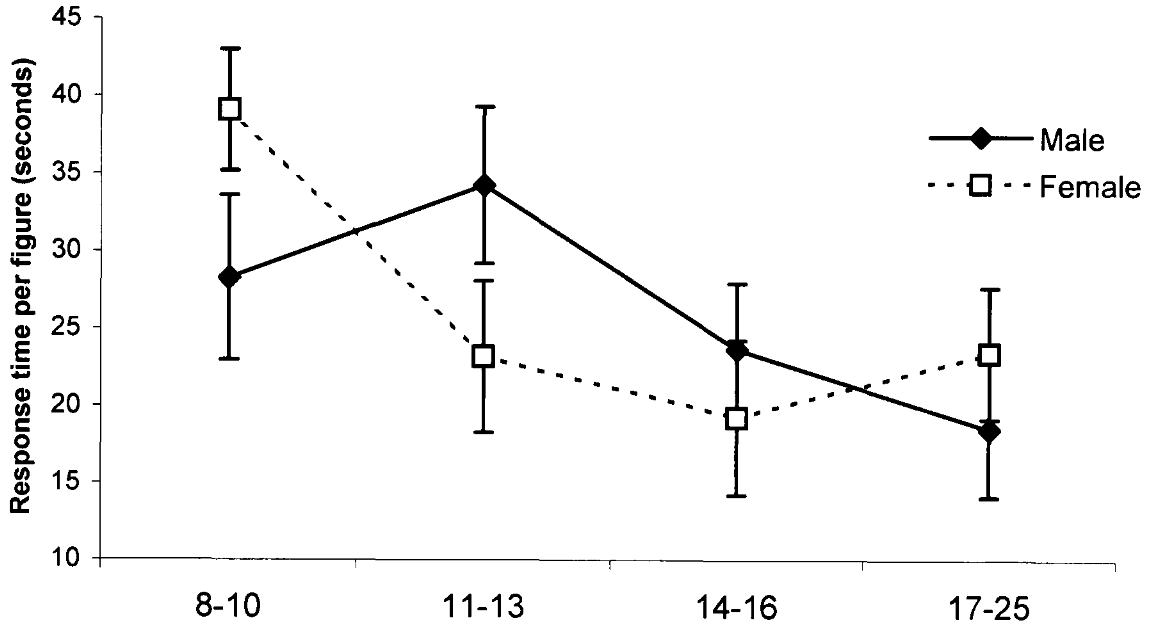


Figure 5.8 Mean time per figure on the EFT for male and female participants in each age group. Error bars show 95% confidence intervals.

Correlations

There was a significant negative correlation between age and time to locate embedded figures in TD sample ($r = -.37$, males $r = -.40$, females $r = -.36$, all $p < .0005$). IQ was also strongly negatively correlated with response time on the EFT across the entire sample (FIQ $r = -.44$, FIQ-BD $r = -.36$, PIQ $r = -.46$, VIQ $r = -.31$; all $p < .0005$), which held for male (FIQ $r = -.47$, FIQ-BD $r = -.41$, PIQ $r = -.51$, VIQ $r = -.35$; all $p < .001$) and female participants (FIQ $r = -.41$, FIQ-BD $r = -.30$, PIQ $r = -.40$, VIQ $r = -.26$; all $p < .007$). Furthermore, EFT and CEFT performance was strongly associated in the TD sample ($r = .50, p < .0005$; males $r = .38, p = .001$, females $r = .60, p < .0005$).

5.3.4 ASD and Control Results

Children's Embedded Figures Test

Three participants (two control, one ASD; 19 to 21 years) were not administered the CEFT because of their age. The exclusion of these participants did not alter the age- and IQ-matching between the ASD and control groups. The distribution of data was examined for the two groups and no outliers were detected. Accuracy data were negatively skewed in both groups (ASD $z = -1.12$, control $z = -3.39$), while the response time data underwent logarithm transformation to approximate normality. Both non-parametric and parametric statistics were used accordingly.

The mean number of correct solutions and time to locate each figure for each group is presented in Table 5.8. One-tailed tests showed that the two groups did not differ significantly on the CEFT in terms of accuracy ($z = 0.57$, $p = .29$) or response time ($t_{(57)} = 0.15$, $p = .44$). No group differences were detected within groups of high IQ (accuracy $z = 0.10$, $p = .49$; response time $t_{(46)} = 0.51$, $p = .31$) or low IQ participants (accuracy $z = 1.33$, $p = .09$; response time $z = 1.28$, $p = .10$).

Correlations

There was no reliable correlation between age or MA and time to locate embedded figures on the CEFT in the ASD group (age $r = -.08$, $p = .67$; MA $r = -.14$, $p = .47$). In contrast, the relationship with age approached statistical significance in the control group ($r = -.34$, $p = .07$) and was highly significant with MA ($r = -.60$, $p = .001$). The strength of the correlation coefficients significantly differed between groups for MA ($z_{r1-r2} = 2.00$, $p = .05$), but not for age ($z_{r1-r2} = 1.00$, $p = .32$).

As found in the TD sample, a strong negative correlation was observed between IQ and response time for the control group (FIQ $r = -.49$, $p = .008$; FIQ-BD $r = -.39$, $p = .04$, PIQ $r = -.57$, $p = .001$; VIQ $r = -.34$, $p = .07$). The dependence of task performance on IQ was not significant in the ASD group (FIQ $r = -.25$, $p = .18$; FIQ-BD $r = -.10$, $p = .60$, PIQ $r = -.26$, $p = .16$; VIQ $r = -.21$, $p = .28$), although the correlation coefficients did not differ between groups (all $z_{r1-r2} < 1.50$, $p > .14$).

Table 5.8 Children's Embedded Figures Test results by group overall, and divided by high and low IQ: Mean (*SD*).

	Group	N	CEFT Number correct (max = 7)	Time per item (seconds)
All	ASD	30	6.3 (0.7)	17.2 (8.2)
	Control	29	5.9 (1.4)	19.3 (12.6)
High IQ	ASD	24	6.3 (0.8)	17.5 (8.7)
	Control	24	6.1 (1.2)	17.2 (11.3)
Low IQ	ASD	6	6.2 (0.8)	16.0 (6.4)
	Control	5	4.8 (1.8)	29.2 (15.4)

Standard Embedded Figures Test

One 15-year-old control participant with MLD experienced repeated failure on the CEFT (located only 3 from 7 figures correctly) and it was considered inappropriate to administer the more difficult EFT. The ASD and control groups remained well matched on age and all IQ measures with the exclusion of this participant.

Table 5.9 presents group means for the number of figures correctly located and the time taken to locate each figure. Accuracy data were negatively skewed in both groups (ASD $z = -4.49$, control $z = -2.32$) and non-parametric analyses were used. Parametric statistics were used on the logarithmic transformed time data. While all ASD participants could correctly locate at least one figure from the EFT, two control participants failed to locate any figures. Significant group differences were found with the ASD group locating more figures than the matched controls ($z = 3.02$, $p = .002$, one-tailed) and being significantly faster ($t_{(59)} = 2.94$, $p = .003$). Group differences were also detected within high-functioning participants (accuracy $z = 2.20$, $p = .01$; response time $t_{(48)} = 2.15$, $p = .02$), and low-functioning participants (accuracy $z = 2.31$, $p = .009$; response time $z = 2.19$, $p = .02$), with the ASD group showing superior performance to the control group.

Table 5.9 Standard Embedded Figures Test results by group overall, and divided by high and low IQ: Mean (*SD*).

	Group	N	EFT Number correct[a] (max = 8)	Time per item[b] (seconds)
All	ASD	31	6.6 (1.5)	24.2 (11.1)
	Control	30	5.1 (2.2)	33.9 (13.6)
High IQ	ASD	25	6.5 (1.6)	24.8 (11.8)
	Control	25	5.4 (2.1)	32.0 (13.1)
Low IQ	ASD	6	7.0 (1.3)	21.9 (8.4)
	Control	5	3.6 (2.1)	43.1 (13.8)

[a]ASD > Control, $p = .002$; ASD_{high} > $Control_{high}$, $p = .01$; ASD_{low} > $Control_{low}$, $p = .009$
[b]ASD < Control, $p = .003$; ASD_{high} < $Control_{high}$ $p = .02$; ASD_{low} < $Control_{low}$, $p = .02$

Correlations

Age did not correlate significantly with response time on the EFT in the ASD ($r = -.01$, $p = .98$) or control group ($r = -.08$, $p = .68$). MA, however, showed a significant negative correlation with response time in the control group ($r = -.49$, $p = .007$), but not in the ASD group ($r = .01$, $p = .94$). The strength of the correlation coefficients were significantly different for ASD and control groups ($z_{r1-r2} = 2.36$, $p = .02$), showing that MA explained significantly more of the variance in response time in the control group (24%) than in the ASD group (0.01%).

A strong negative correlation was found between IQ and time to locate embedded figures in the control group (FIQ $r = -.52$, $p = .003$; PIQ $r = -.66$, $p < .0005$; VIQ $r = -.33$, $p = .07$). This correlation held after excluding the Block Design subtest score from the IQ composite (FIQ-BD $r = -.43$, $p = .02$), which has been reported to correlate with EFT performance (Jarrold et al., 2000; Kalyan-Masih, 1985; Ropar & Mitchell, 2001). In contrast, no IQ measure correlated significantly with EFT performance in the ASD group ($r = -.20$ to $.001$, all $p > .27$). The correlation between PIQ and response time differed reliably in strength between the ASD and control groups ($z_{r1-r2} = 2.22$, $p = .03$).

Interestingly, performance on the CEFT and the EFT was strongly, positively correlated in the control group ($r = .65$, $p < .0005$) but not in the ASD group ($r = .28$, $p = .13$). The difference in strength between the correlation coefficients for the ASD and control groups did not reach statistical significance however ($z_{r1-r2} = 1.81$, $p = .07$). Partial correlations showed that the relationship between the CEFT and EFT remained in the control group after controlling for FIQ ($pr = .53$, $p = .004$), FIQ-BD ($pr = .59$, $p = .001$) and PIQ ($pr = .44$, $p = .02$).

5.3.5 Summary of Embedded Figures Test

In the TD sample, older participants were quicker than younger participants in locating embedded figures, confirming the predicted increase in accuracy with age. IQ also played a significant role in performance, with high-functioning individuals detecting embedded figures with greater speed and accuracy than low-functioning individuals. Males did not differ overall from females, although an interesting age group by gender interaction was found, suggesting females improve in EFT performance at an earlier age than males. This finding did not confirm the suggestion that male superiority appears after 14 years of age (Voyer et al., 1995).

ASD participants were more accurate and significantly quicker than their matched controls at locating embedded figures on the standard EFT version of the task, but not the CEFT. This advantage was evident in high-functioning as well as low-functioning participants. Moreover, the low-functioning ASD participants performed at an equivalent level to 17-25-year-old TD participants on the EFT (cf. Table 5.7 and Table 5.9). The finding of ASD superiority when the stimuli consisted of non-meaningful geometric forms (EFT), but not meaningful pictures (CEFT), is counter to the prediction that meaningfulness may induce a stronger gestalt. However, it is possible that ceiling effects on the CEFT obscured any group differences.

Unlike the TD and control participants, age and IQ was not related to test performance in the ASD group. It is notable that the low and high IQ ASD groups differed very little on the CEFT and EFT. Furthermore, CEFT and EFT performance did not correlate in the ASD group, in contrast to the TD and control groups. This may lead to the suggestion that different processes may be operating for meaningful versus non-meaningful contexts for individuals with ASD. However, the variance of the data was slightly smaller in the ASD group than the control group, which may have limited the opportunity for significant correlations in the ASD sample.

As a significant difference between ASD and control groups was detected on the EFT (and not the CEFT), the mean time to locate figures on the standard version of the task was taken as the main dependent variable and used in subsequent analyses. This variable incorporates both accuracy and speed of detection as maximum time is given to items not found. In categorical analyses, individuals were divided by whether their mean response time on the EFT was less than, or greater than, 20 seconds (indicating weak versus strong coherence respectively).

5.4 Impossible Figures Task

Impossible figures are drawings of geometric forms that would be impossible to construct in three-dimensions. A classic example is the Penrose triangle, first described in 1958 (see Figure 5.9). The triangle is geometrically possible at each corner, but a contradiction is apparent when viewing the triangle as a whole. The figure is therefore locally possible but globally impossible when a three-dimensional interpretation is attempted. Impossibility is said to be an emergent property of the whole figure, which requires the integration of parts in order to be detected (Young & Deregowski, 1981). As visual coherence is required to identify impossibility, an impossible figures detection task was included as a measure of configural processing at a low visuo-spatial level.

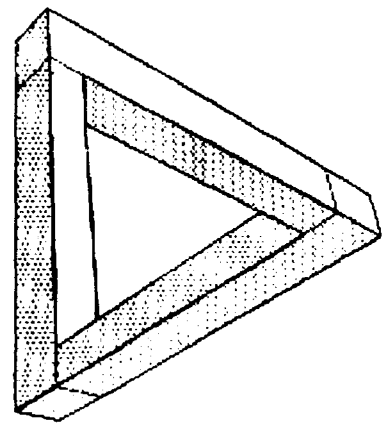


Figure 5.9 An impossible triangle (Penrose & Penrose, 1958).

Detecting geometric impossibility in ASD

Studies within the autism literature have shown that individuals with ASD do not readily perceive the impossibility of impossible figures. Mottron and Belleville (1993) reported impairment in the perception of impossible figures in their case study of EC, a 34-year-old male with Asperger syndrome and exceptional graphic drawing skills. Several areas of EC's perceptual processing were intact, for example he was able to use monocular cues in depth perception and was sensitive to visual illusions. However EC was less efficient than controls in detecting the geometrical impossibility of figures when briefly presented (100 to 800 ms), but could detect impossibility if viewed for a longer time (15 seconds). When asked to draw impossible figures from memory, E.C. would produce globally coherent figures that did not include the impossible element. Again, this was only evident when he viewed the figures for short durations; E.C. could correctly draw the impossible element when the figure was exposed for longer.

Mottron, Belleville and Menard (1999) explored the perception of impossibility in 10 high-functioning adults and adolescents with autism and 10 age- and IQ-matched controls. They found that both groups took longer to copy impossible figures compared to their possible counterparts, even though the figures were matched for number and type of features. The difference between conditions was less apparent in the autism group however, suggesting that these individuals might not be as affected by geometric impossibility.

Rodgers (2000) also examined the ability to detect geometric impossibility in her study of eight adults with Asperger syndrome. Pairs of matching possible and impossible figures were presented simultaneously and participants were asked to indicate which of the two figures was impossible. She found that the individuals with Asperger syndrome made significantly more errors in identifying the impossible figure compared to an age- and IQ-matched control group.

Detecting geometric impossibility in typical development

Impossible figure stimuli have been used in several studies of human perception and object recognition (e.g., speed of detection, Donnelly, Found, & Muller, 1999; implicit memory tasks, Schacter, Cooper, & Delaney, 1990; P. Williams & Tarr, 1997; computational modelling for detection, Cowie & Perrott, 1993). Only a few studies have looked at developmental effects in the ability to detect impossibility (e.g., Deregowski, 1969; Young & Deregowski, 1981).

Deregowski (1969) found that young children's perception of impossibility was directly related to their ability to observe the three-dimensional aspect of line drawings. Nine to 10-year-old children were asked to copy possible and impossible forms of the trident figure (see Figure 5.10). Children categorised as *three-dimensional perceivers* spent more time looking at the impossible trident than at the possible trident, whereas *two-dimensional perceivers* did not differ significantly in the time spent viewing each figure. The perceptual inconsistencies within the impossible figure did not impede the task of copying in the two-dimensional perceivers. An alternative interpretation might characterise the so-called two-dimensional perceivers as being less aware of the whole form and more local in processing style.

In a later study, Young and Deregowski (1981) investigated the extent to which children (7 to 14 years) perceive the impossibility of geometric figures. The ability to identify the impossible figure from a pair of corresponding possible and impossible figures was found to improve with age. Seven-year-olds did not readily perceive impossibility as they took the same amount of time to view impossible and possible figures in order to draw them from memory. Ten-year-olds, in contrast, spent longer looking at the impossible than possible form. Contrary to the findings from Deregowski (1969), 7-year-olds were familiar with the conventions used to represent three-dimensional objects in drawings and this did not explain their failure to detect impossibility. The authors concluded the younger children lacked the ability to construct a mental representation of the object, including the interrelationships between parts, which was necessary for the perception of impossibility.

(a) (b)

Figure 5.10 Examples of the possible (a) and impossible (b) forms of the trident figure.

In light of the findings described above, detecting geometric impossibility was included as a measure of possible individual differences in central coherence. In the Impossible Figures Task, developed for the present study, participants were assessed on their speed and accuracy in determining whether individually presented geometric forms were possible or impossible.

Predictions

From the findings of Young and Deregowski (1981), it was predicted that ability to detect impossibility of geometric forms would increase with age in TD participants. As gender effects have not been examined in the literature on impossible figure detection, no predictions of gender differences were made. In line with the finding that individuals with ASD have shown problems in detecting geometric impossibility (Mottron & Belleville, 1993; Mottron, Belleville et al., 1999; Rodgers, 2000), it was predicted that the ASD group would be less proficient at detecting global possibility of both possible and impossible geometric figures than age- and IQ-matched controls. In addition, it was anticipated that individual differences would be found in TD and ASD, over and above the effects of age or IQ.

5.4.2 Method

Materials

Test stimuli consisted of 20 geometric figures (10 possible and 10 impossible figures, in matched pairs), adapted from Young and Deregowski (1981), Robinson and Wilson (1973) and Térouanne (1980). The figures (shown approximately 40 percent of their actual size) are presented in Figure 5.11.

Task Development

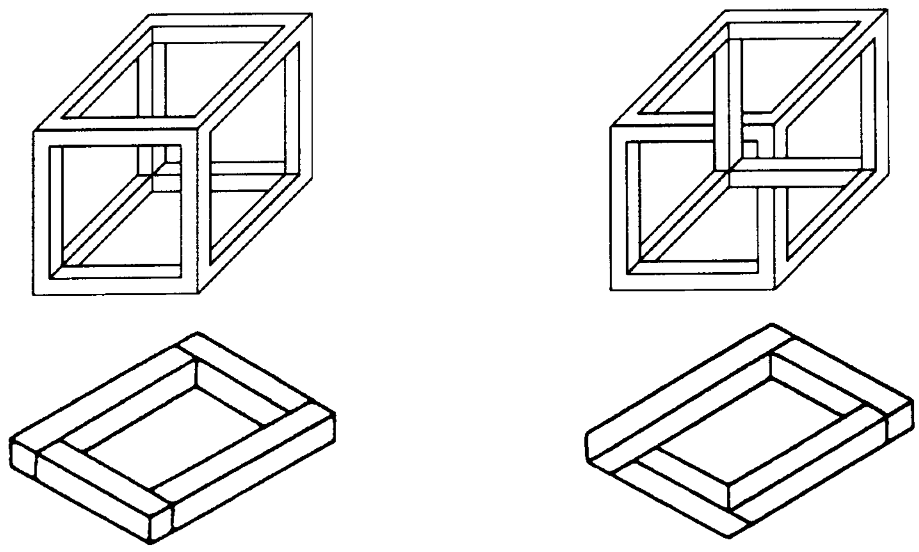
The Impossible Figures Task was piloted on 32 TD participants (8 to 15 years). Participants were presented with a sequence of geometric forms and asked whether the object was possible or impossible. After the first wave of piloting it was apparent that the concept of impossibility was difficult for young children to comprehend, especially when the figures were presented in isolation. The introduction phase of the task was therefore developed to simultaneously show possible and impossible forms of the same figure. Once the participant could reliably distinguish between basic possible and impossible forms, the task progressed to identifying whether individual figures were possible or impossible.

The pilot sample showed a range of ability in detecting geometric impossibility. The average rate of detection across the set of 20 figures was 74% (SD = 17), with 71% of the possible figures identified correctly and 78% of the impossible figures. As an adequate range of difficulty was found in the pilot study, with no ceiling or floor effects, the full set of 20 impossible and possible figures was used in the main study.

Item **Possible Form** **Impossible Form**

Crate

Diamond



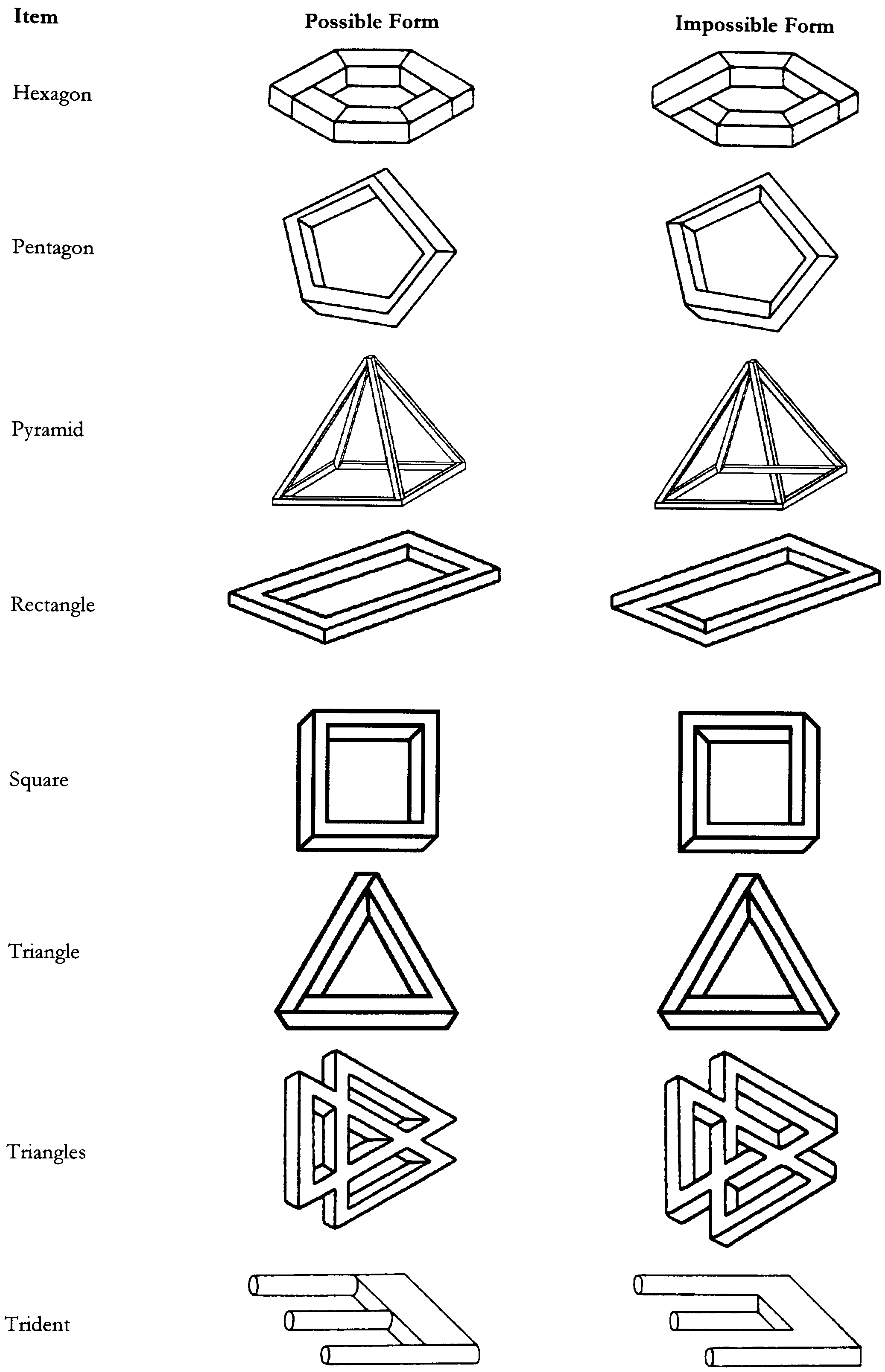


Figure 5.11 Possible and impossible stimuli used in the Impossible Figures Task.

Procedure

The task began with an introduction phase to ensure that participants understood the concept of geometrical possibility. Possible and impossible forms of the simple trident were presented together on the touch screen (see Figure 5.12 a and b). The researcher told the participant that one of the two objects was real and could be made out of wood, but the other could not because there was something wrong with the drawing. The participant was then asked to select which object was real or possible. If the participant touched the possible trident a congratulatory sound was played (Windows sound file: utopia asterisk.wav). If the participant touched the impossible trident the researcher explained how the figure was not possible and encouraged the participant to touch the possible figure. Once the correct selection was made a second example appeared consisting of the possible and impossible forms of a rod (see Figure 5.12 c and d). The participant was again encouraged to select the possible figure and the program moved on only after the correct selection was made.

Four practice trials were administered following the introduction. Participants were told figures would appear one at a time, and they had to decide whether each one was possible or impossible. Participants indicated their answer by touching the word *possible* or *impossible* presented at the bottom left and right of the screen respectively. The four stimuli from the introduction phase were used in the practice trials. Corrective feedback was provided during the practice trials but not during the test phase.

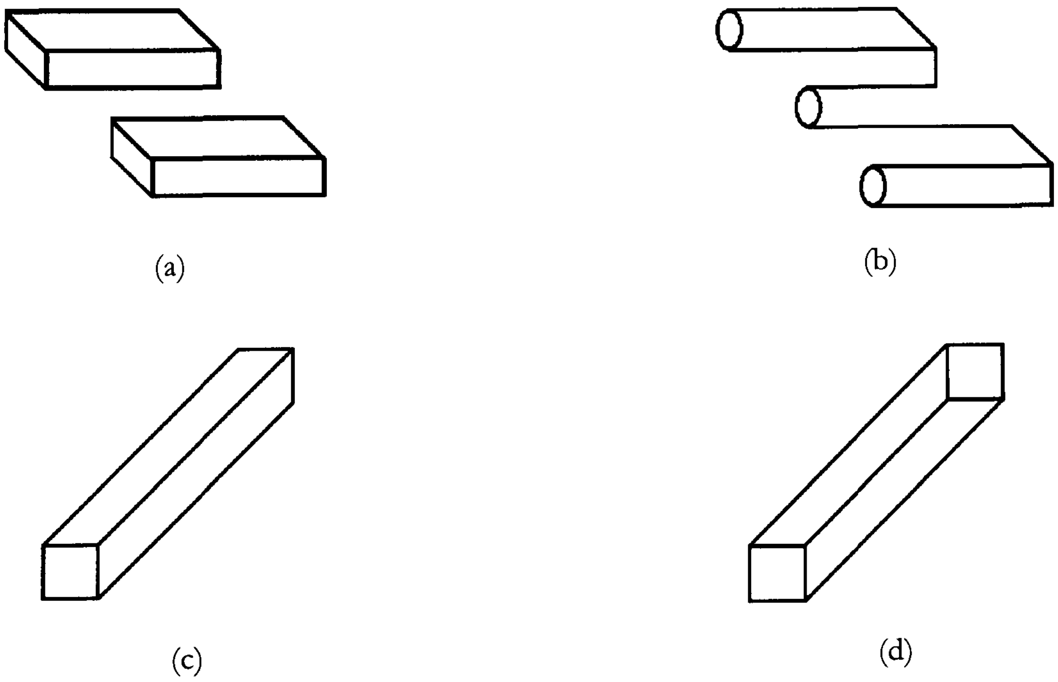


Figure 5.12 Practice items: simple trident in (a) possible and (b) impossible forms, and rod in (c) possible and (d) impossible forms (shown 50% of actual stimulus size).

When it was established that the participant understood the task requirements, the test phase began. Participants were reminded that they were being timed by the computer so to make their decision quickly, but as accurately as possible. The set of 20 possible and impossible figures was presented in a fixed random order. Each figure remained on the screen until the participant made a response. Accuracy and time to respond from the onset of each figure were recorded.

5.4.3 *Typical Development Results*

All TD participants were administered the Impossible Figures Task, although data files were lost for two females (from 11-13 and 14-16 year groups) and these participants were excluded from analysis. All participants demonstrated an understanding of geometric impossibility from the introduction and practice trials.

As participants had an unlimited time to respond to each figure, response times for individual items were inspected for outliers. There were 15 occasions where response time exceeded 10 seconds (9 for impossible figures, 6 for possible figures), with 8 of these originating from one 9-year-old female (range: 12 to 22 seconds). All response times were consequently capped to 10 seconds in order to eliminate the effect of extreme values. There were four occasions where response time was less than 200 ms (range 50 to 105 ms, all originating from different participants) suggesting the trial moved on prematurely, before the participant made a valid response. These items were treated as missing data and response time and accuracy were based on the number of valid items for each participant.

Item analysis

The proportions of figures correctly identified as possible or impossible for each individual figure in each age group and gender are presented in Appendix G. Detection rates were variable between figures and confirmed the finding of Young and Deregowski (1981) that, regardless of age, the impossibility of some types of figures was relatively easy or difficult to detect. Both possible and impossible forms of the triangle and pentagon were most difficult to judge for all participants (correct detection rates: impossible triangle 60%, possible triangle 67%, impossible pentagon 66%, possible pentagon 69%). These two items were removed from analysis as it was considered that item difficulty might obscure individual differences in detecting impossibility.

Accuracy

Table 5.10 presents the mean proportion of correctly judged possible and impossible figures and the corresponding nonparametric index of sensitivity[1] A' (combining correct detections of impossible figures and incorrect detections of possible figures) for each age group. As the data were negatively skewed in each age group (z-scores -0.8 to -10.5),

---

[1] The values of A' range from 0 to 1, with a value of 1 indicating perfect discrimination and 0.5 indicating chance performance (an A' of 0.0 would indicate perfectly inaccurate performance). $A' = 0.5 + [(H - FA)(1 + H - FA)] / [4(H)(1 - FA)]$ if the hit rate (H) is greater than the false alarm rate (FA), or $A' = 0.5 - [(FA - H)(1 + FA - H)] / [4(FA)(1 - H)]$ otherwise (Grier, 1971).

non-parametric statistical tests were used. The non-parametric measure of bias[2], B'', was also calculated and no age group differences ($\chi^2 = 6.32$, $p = .10$) or gender effects ($z = 0.69$, $p = .49$) were detected. Mean B'' values were positive for all age groups (range .17 to .47) indicating a general bias towards responding "possible".

Two male participants (14 years, 20 years) performed below chance on A' (scores of .35 and .29 respectively) and their reaction time and bias data were examined. As there was evidence that each participant was not performing the task (e.g., extremely fast response rates), a conservative approach was taken and their data were removed from analyses.

Table 5.10 Proportion of correctly judged possible and impossible figures and the corresponding measure of sensitivity, A', for each age group: Mean (*SD*)

Age group (years)	N	Possible Figures[a]	Impossible Figures[b]	A'[c]
8 - 10	54	0.80 (0.20)	0.75 (0.18)	0.84 (0.13)
11 - 13	42	0.90 (0.15)	0.83 (0.15)	0.92 (0.08)
14 - 16	49	0.96 (0.11)	0.86 (0.16)	0.95 (0.07)
17 - 25	55	0.94 (0.13)	0.86 (0.17)	0.94 (0.10)

[a] 8-10 < 11-13, 14-16, 17-25; $p < .008$; 11-13 < 14-16; $p = .007$
[b] 8-10 < 11-13, 14-16, 17-25; $p < .02$
[c] 8-10 < 11-13, 14-16, 17-25; $p < .003$; 11-13 < 14-16, 17-25; $p < .03$

As A' incorporated correct detections of both possible and impossible figures, this index is reported in the subsequent analyses. Significant effects of age group were found on A' ($\chi^2 = 35.50$, $p < .0005$). Participants in the youngest age group were less able to discriminate between possible and impossible figures than the three older age groups (all $z > 3.04$, $p < .003$). The 11-13 year group was also less able to discriminate the two figure types relative to the two oldest age groups (all $z > 2.29$, $p < .03$).

Table 5.11 presents mean detection rates and A' values for male and female participants in each age group. A main effect of gender was found with males distinguishing between possible and impossible figures with greater accuracy than females ($z = 3.35$, $p = .002$). When analysing by age group, this gender effect was significant in the oldest age group ($z = 2.95$, $p = .003$), with a trend shown in the youngest age group ($z = 1.88$, $p = .06$).

---

2 As described by Donaldson (1992), the non-parametric model using the B'' measure of bias appears to be as or more robust than other more widely used signal detection models (i.e., B', Grier, 1971). The values of B'' range from −1.0 to +1.0, with 0.0 indicating no bias, positive values indicating conservative bias (i.e., greater proportion of omissions over false alarms) and negative values indicating a liberal bias (i.e., greater proportion of false alarms over omissions). B'' = [(1-H)(1-FA) − (H)(FA)] / [(1-H)(1-FA) + (H)(FA)].

Table 5.11 Proportion of correctly judged possible and impossible figures and the corresponding measure of sensitivity, A', for males and females in each age group and overall: *M* (*SD*)

Age group (years)		Possible Figures[a]	Impossible Figures[b]	A'[c]
8 - 10	*male*	0.87 (0.18)	0.79 (0.12)	0.89 (0.10)
	female	0.76 (0.21)	0.73 (0.21)	0.81 (0.14)
11 - 13	*male*	0.92 (0.10)	0.87 (0.12)	0.94 (0.05)
	female	0.89 (0.19)	0.79 (0.17)	0.90 (0.10)
14 - 16	*male*	0.95 (0.10)	0.84 (0.18)	0.94 (0.08)
	female	0.97 (0.11)	0.88 (0.13)	0.96 (0.07)
17 - 25	*male*	0.99 (0.04)	0.90 (0.16)	0.97 (0.05)
	female	0.90 (0.17)	0.82 (0.16)	0.90 (0.12)
All	*male*	0.94 (0.12)	0.85 (0.16)	0.94 (0.08)
	female	0.86 (0.19)	0.79 (0.18)	0.86 (0.19)

[a]males: 8-10, 11-13 < 17-25; $p < .02$; females: 8-10 < 11-13, 14-16, 17-25; $p < .01$; 11-13, 17-25 < 14-16; $p < .05$
[b] males: 8-10 < 11-13, 14-16, 17-25; $p < .05$; females: 8-10 < 17-25; $p = .006$
[c] males: 8-10 < 11-13, 14-16, 17-25; $p < .05$; 11-13 < 17-25, $p = .06$
females: 8-10 < 11-13, 14-16, 17-25; $p < .05$; 11-13 < 14-16; $p = .007$

Mean A' values for male and female participants are plotted as a function of age group in Figure 5.13. It appears that effects of age group on A' are greater for females than males (i.e., steeper graph for females than males), although significant effects of age group were found for both genders (males $\chi^2 = 14.58$, $p = .002$; females $\chi^2 = 20.77$, $p < .0005$). The pattern of results shows the possible contribution of general ability to task performance, as male superiority in FIQ is apparent in the youngest and oldest age groups. Furthermore, the switch in performance for male and females at 14-16 years may be related to the higher VIQ in females relative to males in this age group, however the non-parametric statistical analyses do not allow for the covariation of IQ to establish this association.

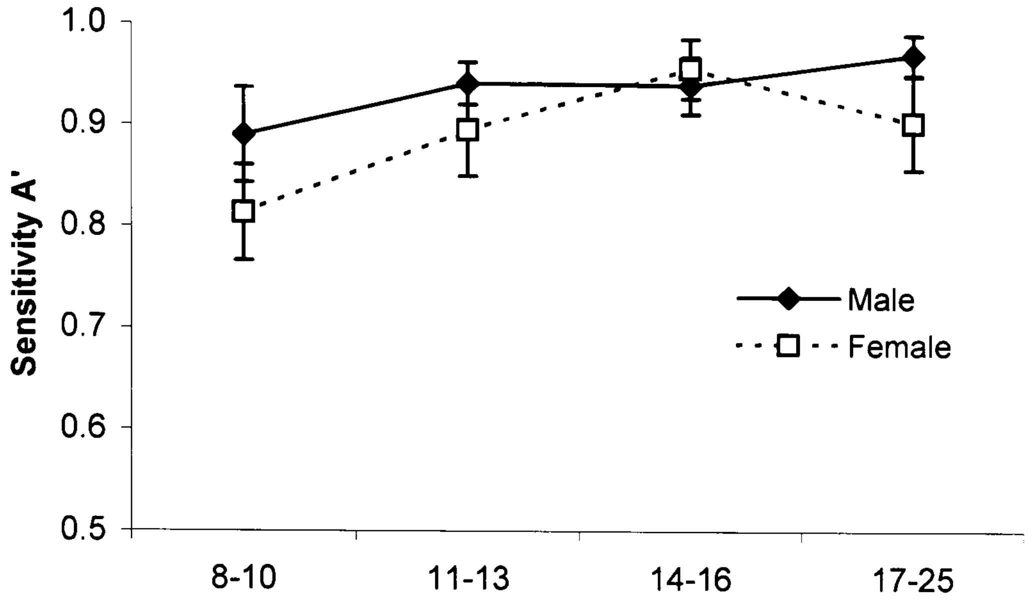


Figure 5.13 Mean sensitivity indices for detecting geometric impossibility for male and female TD participants in each age group. Error bars show 95% confidence intervals.

Response time

There was evidence of a speed-accuracy trade-off in the youngest age group with A' showing a significant positive correlation to mean response time (r_s = .29, p = .03). The relationship between response time and accuracy by figure type did not reach significance in this group however (possible r_s = .19, p = .17; impossible r_s = .24, p = .08). Negative (and non-significant) correlations between response time and accuracy were observed in all older age groups (r_s = -.05 to -.25, all p > .10).

Time for correct responses were compared between age groups (see Table 5.12). Analyses were also conducted on all times (i.e., including both correct and incorrect responses), but as the pattern of results did not change, correct times are reported. Mean response times were positively skewed in each age group (z-scores 0.6 to 10.1), and non-parametric analyses were used as a consequence. No age group effects were found on the mean time to correctly detect impossible figures (χ^2 = 2.43, p = .49), or possible figures (χ^2 = 7.17, p = .07). No effect of gender was found on mean correct response time in the entire sample (all z < 1.22, p > .21), or within each age group (all z < 1.75, p > .07).

Table 5.12 Response time (in seconds) for correct detections of possible figures and impossible figures for each age group: Mean (*SD*)

Age group (years)	N	Possible Figures	Impossible Figures
8 - 10	54	2.01 (0.91)	2.39 (0.99)
11 - 13	42	2.13 (0.78)	2.41 (0.77)
14 - 16	49	1.75 (0.47)	2.22 (0.94)
17 - 25	55	1.85 (0.56)	2.43 (0.90)

An interesting observation was that individuals appeared to spend longer judging impossible figures compared to possible figures. Wilcoxon's Signed Ranks tests confirmed this, showing that participants within all age groups took longer to respond to impossible figures than possible figures (all $p < .0005$). At an individual level, 76% of participants in the 8-10 year group (41 from 54) had longer response times for impossible than possible figures. This was true for 74% of the 11-13 year group (31 from 42), 80% of the 14-16 year group (39 from 49) and 85% of the 17-25 year group (47 from 55). The subgroup of individuals who look longer to respond to possible figures than impossible did not differ in age, IQ or gender to the majority who took longer to judge impossible figures relative to possible figures.

Correlations

A significant positive correlation between A' and age was found in the TD sample ($r_s = .37$, males $r_s = .45$, females $r_s = .30$, all $p < .003$). IQ was also strongly related to the correct detection of possible and impossible figures in the entire sample (FIQ $r_s = .50$; FIQ-BD $r_s = .42$; VIQ $r_s = .38$; PIQ $r_s = .50$), in males (FIQ $r_s = .43$; FIQ-BD $r_s = .36$; VIQ $r_s = .32$; PIQ $r_s = .50$), and in females (FIQ $r_s = .53$; FIQ-BD $r_s = .44$; VIQ $r_s = .40$; PIQ $r_s = .47$, all $p < .0005$).

5.4.4 *ASD and Control Results*

One ASD participant was not administered the Impossible Figures Task and the data file was lost for one control participant. The groups remained well matched on age and IQ measures after both participants were excluded from analyses. All ASD and control participants demonstrated an understanding of geometric impossibility from the introduction and practice phase of the task.

Item analysis

The proportions of figures correctly identified as possible or impossible for each individual figure in the ASD and control group are presented in Appendix G. In line with the TD analyses, the two most difficult figures to categorise in both possible and impossible forms (triangle and pentagon) were removed from analysis. Mean accuracy and response times are therefore based on 8 possible and 8 impossible figures.

Response time data were inspected for outliers in the ASD and control groups. No time data were less than 250 ms, while on 4 occasions response time exceeded 10 seconds (3 for impossible figures, 1 for possible figure). Two extreme times were from one ASD participant (12 to 27 seconds), while the remaining times came from control participants (11 to 12 seconds). All response times were consequently capped to 10 seconds.

Accuracy

Table 5.13 presents the proportion of correctly judged possible and impossible figures for each group and the signal detection index of sensitivity, A'. As the data were strongly negatively skewed (z-scores -2.2 to -4.0) non-parametric analyses were conducted. Two control participants and one ASD participant (all FIQ < 74) performed below chance on A' and an examination of their task performance (reaction time and bias) was made. As it could not be ascertained whether these participants found the task difficult, rather than misunderstood the instructions, their data were removed from analyses.

Table 5.13 Proportion of correctly judged possible and impossible figures and the corresponding measure of sensitivity, A', for each group: Mean (*SD*)

	Group	N	Possible Figures	Impossible Figures[b]	A'[a]
All	ASD	29	0.85 (0.18)	0.75 (0.19)	0.85 (0.14)
	Control	28	0.91 (0.13)	0.80 (0.22)	0.90 (0.12)
High IQ	ASD	24	0.90 (0.13)	0.76 (0.20)	0.88 (0.13)
	Control	25	0.92 (0.12)	0.85 (0.17)	0.93 (0.08)
Low IQ	ASD	5	0.60 (0.19)	0.70 (0.14)	0.72 (0.11)
	Control	3	0.79 (0.14)	0.46 (0.31)	0.66 (0.20)

[a]ASD < Control, p = .05
[b]$\text{ASD}_{\text{high}} < \text{Control}_{\text{high}}$ p = .04

ASD and control groups did not differ on non-parametric measure of bias, B'' (z = 0.11, p = .92). No group differences were detected on the mean number of correct judgments of possible figures (z = 1.34, p = .09) and impossible figures (z = 1.46, p = .07), using one-tailed Mann-Whitney tests. On the overall measure of sensitivity (A') however, ASD participants were significantly less able to discriminate between possible and impossible figures than the control group (z = 1.68, p = .05). When excluding participants with low IQ from analyses, group differences failed to emerge on A' to a significant level

($z = 1.60$, $p = .06$). Comparing ASD and control participants with low IQ also did not reveal any significant group differences ($z = 0.60, p = .29$).

Response time

In the ASD group a significant positive correlation was found between A' and mean overall reaction time ($r_s = .43, p = .02$). A strong association was found between response time and accuracy for possible figures ($r_s = .45$, $p = .02$), but not for impossible figures ($r_s = -.18, p = .35$). William's equation (1959; as cited in Howell, 2002, p. 281) to test the difference between two non-independent correlations found these two coefficients to be significantly different in the ASD group ($t_{(26)} = 7.01$, $p < .0005$). No such speed-accuracy trade off was found in the control group (all $r_s = -.21$ to $-.08, p > .26$).

Mean response times for correct responses in detecting possible and impossible figures are presented in Table 5.14 for each group. ASD and control groups did not differ in the mean time to correctly identify possible figures ($z = 0.11$, $p = .46$) or impossible figures ($z = 0.19, p = .42$). No group differences were detected within groups of high IQ (possible $z = 0.28$, $p = .39$; impossible $z = 0.42$, $p = .34$) or low IQ participants (possible $z = 1.04$, $p = .15$; impossible $z = 1.04, p = .15$).

As found in the TD sample, ASD and control participants took longer to identify impossible figures correctly than possible figures (Wilcoxon's Signed Ranks Tests, all $p < .02$). At an individual level, 79% of the ASD group (23 from 29) had longer response times for impossible than possible figures. This was apparent for 71% of the control group (20 from 28). Within group comparisons showed that individuals who took longer to respond to impossible figures compared to possible figures were of similar age and IQ to those with the reverse pattern (ASD all $z < 0.86, p < .38$; control $z < 1.56, p > .11$).

Table 5.14 Response time (in seconds) for correct detections of possible figures and impossible figures for each group: Mean (*SD*)

	Group	N	Possible Figures	Impossible Figures
All	ASD	29	1.81 (0.56)	2.15 (0.74)
	Control	28	1.90 (0.76)	2.18 (0.71)
High IQ	ASD	24	1.85 (0.59)	2.25 (0.76)
	Control	25	1.86 (0.72)	2.17 (0.70)
Low IQ	ASD	5	1.62 (0.40)	1.65 (0.39)
	Control	3	2.27 (1.17)	2.19 (0.99)

Correlations

The sensitivity measure A' did not correlate with age in the ASD ($r_s = .17$, $p = .37$) or control group ($r_s = -.11$, $p = .58$), although a significant positive correlation was found between MA and A' in both groups (ASD $r_s = .46$, $p = .01$; control $r_s = .52$, $p = .005$). A strong positive relationship was also found between IQ and A' in the control group (FIQ $r_s = .72$, FIQ-BD $r_s = .70$, VIQ $r_s = .65$, PIQ $r_s = .61$; all $p < .0005$). The strength of the correlations between IQ and A' were lower in the ASD group (FIQ $r_s = .46$, FIQ-BD $r_s = .42$, PIQ $r_s = .42$; all $p < .03$; not significant for VIQ $r_s = .29$, $p = .13$), but did not differ significantly from the control group (all $z_{r1-r2} < 1.70$, $p > .08$).

5.4.5 *Summary of Impossible Figures Task*

The ability to discriminate between possible and impossible geometric figures was found to increase with age and ability in TD. In line with predictions from weak central coherence theory, ASD participants were impaired in their ability to determine geometric impossibility compared to age- and IQ-matched controls.

Previous studies have shown that the perception of impossibility is indicated by longer looking times to impossible compared to possible forms when asked to draw the figure (Deregowski, 1969; Young & Deregowski, 1981). A comparable finding was found in the present study, whereby the majority of participants in all groups took longer to judge impossible figures correctly than possible figures. This suggests that the incongruity of an impossible figure is not an immediate emergent property, but perhaps perceived through a process of integration or serial search. In contrast, the global coherence of possible figures may have a "pop out" effect.

A further interesting finding was the strong association between accuracy and response time in ASD participants for determining possible figures, but not impossible figures. This speed-accuracy trade-off was observed in the youngest TD age group, but did not differentiate between possible/impossible figure type. This suggests that individuals with ASD may be performing an exhaustive serial search on possible figures, with the implication that perception of global coherence has less immediacy in individuals with ASD compared to TD.

The measure of sensitivity in discriminating between possible and impossible figures (A') was taken as the key variable from the Impossible Figures Task. Relatively poor ability to determine global geometric impossibility (suggesting weak central coherence) was taken as A' values less than .80. This cut-off was used in further categorical analyses.

5.5 Navon Similarity Judgment Task

One of the most extensively used tasks to study visual perception at local and global levels consists of hierarchical figures (see Figure 5.14 for an example). As first described by Navon (1977), hierarchical figures can be analysed at both the global (i.e., the overall shape of the stimulus) and the local level (i.e., the individual features that comprise the overall shape). According to classic research begun by Navon (1977; reviewed by Navon, 2003), higher level or global information is processed faster and more accurately than lower level or local information, an effect known as *global advantage*. Furthermore, identifying the global form can be achieved without interference from the local level, but it is often difficult to ignore the global form when processing the local level (the *global-to-local interference* effect). Both phenomena make up the *global-precedence hypothesis* (Kimchi, 1992) and are often demonstrated in typical-developing populations (see Chapter 3, Section 3.2, for a review).

```
H       H
H H     H
H   H   H
H     H H
H       H
```

Figure 5.14 Example of a hierarchical figure: a global letter (N) composed of local letters (H).

Hierarchical figure perception in ASD

In accordance with weak central coherence theory it is predicted that individuals with ASD will show precedence towards processing local information over global information when viewing hierarchical figures. It has been shown that individuals with autism demonstrate a lack of global advantage, although results are mixed (Mottron & Belleville, 1993; Mottron, Burack et al., 1999; Ozonoff et al., 1994; Rinehart et al., 2000; Rodgers, 2000; see Chapter 2 for a review). It has been documented that the classic Navon task is very sensitive to methodological variation (Kimchi, 1992), which may explain the inconsistent results: for example, whether participants are directed toward attending to the local or global level (selective attention tasks) or if the target appears randomly at either level (divided attention tasks). In an undirected (divided attention) situation, the preferred processing level appears to be the local level in individuals with autism (Plaisted et al., 1999).

Hierarchical figure perception in typical development

Age-related changes have been found in TD populations on tasks using Navon hierarchical figures (see Chapter 3 for a review). Kramer et al. (1996) for example, found that older children (7 to 12 years) were more perceptually biased toward the global form than younger children (4 to 6 years). They also found that boys had a greater tendency towards the global form than females at all ages. Dukette and Stiles (1996) also reported an increase in global bias with development (comparing children aged 4 years, 6 years, and adults), even when the perceptual salience of the form had been degraded. In contrast to Kramer et al., females were more biased towards the global form than males, although this gender effect was apparent when the hierarchical figures were composed of letters, but not geometric forms. Kimchi (1990) found a similar global bias and sensitivity to changes to stimulus density in children (3 to 9 years) as typically found in adults.

Task Development

As an initial step to assess the perception of hierarchical figures in the present study, a divided attention task was designed and piloted on 37 TD children (7 to 15 years). The participant was required to search for a given letter in a hierarchical figure that could appear either at the local or global level. Six figures (Nh, Na, Hn, Ha, An, Ah)[3] were presented in six densities (comprising 4, 5, 6, 7, 8, 9 local elements in width and height) and in three conditions (global target, local target, target absent) resulting in 108 trials (presented in two blocks of 54 trials). Due to the high number of trials required and great within-subject variability found, the task was abandoned in preference to using the hierarchical figures in a forced-choice similarity judgment task.

In the revised task participants were briefly shown a standard hierarchical figure and were asked to select from two comparison figures the one that looked most like the standard figure. One comparison figure was composed of the same local elements but formed another global letter: the *local* match. The other comparison figure was composed of different local elements but formed the same global letter as the standard: the *global* match. This task design was deemed more suitable for measuring individual differences in local-global processing style across a wide age and ability span than divided or selective attention tasks. First, as the task is non-directive, participants' natural tendency or preference towards attending to the global or local levels (or a combination of both) can be assessed. Secondly, as there are no right or wrong answers the task is non-threatening and requires a minimal response of pointing to or touching a comparison stimulus. Finally, the task assesses response selection rather than speed of processing which is typically impaired in low-functioning individuals (Anderson, 2001).

---

3 The upper-case letter denotes the global form and the lower-case letter denotes the local element.

Exposure duration

The similarity-judgment paradigm has proved successful in testing spatial cognition in young children as well as adults (e.g., Dukette & Stiles, 1996; Kimchi, 1990; Kimchi & Palmer, 1982; Kramer et al., 1996). In the typical task the standard figure is presented at the same time as the two comparison figures and is left in view until a choice is made. Participants are asked to give their first, most immediate impression of a match, although they have unlimited time to make their selection. This unrestricted exposure to the standard figure can mean that both local and global levels are analysed. For example, Brosnan et al. (2004) reported that most of their sample of low-functioning children with autism (79%) could identify both levels of hierarchical stimuli that were presented with unlimited time, with 80% naming the global letter first. Exposure duration is consequently an important factor that can affect global advantage (Kimchi, 1992). Jolliffe and Baron-Cohen (1997) suggested that the failure of Ozonoff et al. (1994) to detect differences in their autism sample on the Navon task was because the exposure time used (1000 ms) was too long to produce this subtle effect. Mottron and Belleville (1993) found an absence of global precedence in a man with Asperger syndrome, but only when stimuli appeared very quickly (10 to 25 ms). In order to measure participants' initial impression of the standard figure, a presentation time of 250 ms was used in the present study.

Density of local elements

The spacing of the local elements is a factor that is known to affect global advantage (Kimchi, 1992; Kimchi & Palmer, 1982). Typically, more sparse stimuli produce local advantage, while more dense stimuli result in global advantage (Martin, 1979). Kimchi suggested that greater spacing violates the "goodness" of the global form, which makes it more difficult to "extract the identity of the global letter than the identity of the local letter" (p. 28). Conversely, an increased number of local elements can result in the elements being perceived as texture and the global form becomes relatively more salient. To assess participants' sensitivity to the number of local elements, the hierarchical figures in the present study were shown at three levels of density.

Predictions

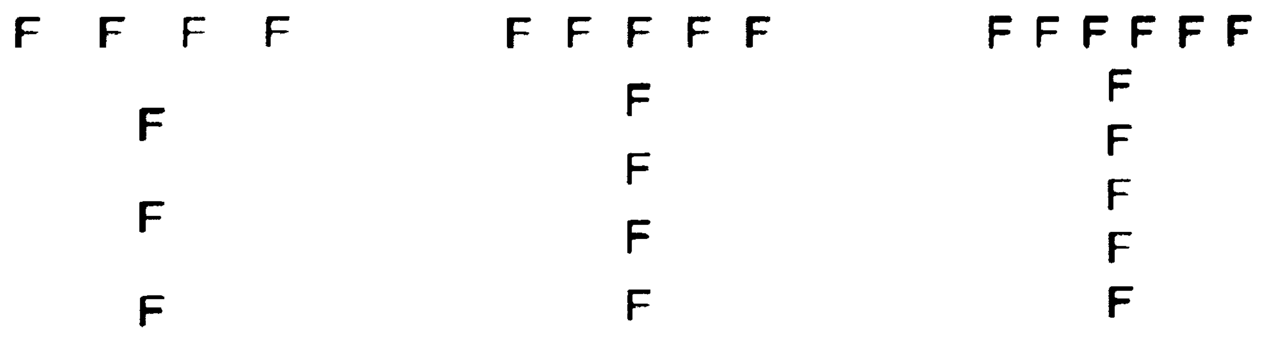
It was predicted that TD participants would be biased towards making similarity judgments based on the global form over the local elements, with a stronger global bias shown in older participants (Kramer et al., 1996). Similarly, with a greater number of local elements, it was predicted that the global form would be more salient than local elements for all participants, but older participants would be more sensitive to this change in density (Dukette & Stiles, 1996; Kimchi, 1990; Stiles et al., 1991).

As gender effects have shown mixed results in the literature, no a priori predictions were made. Studies employing hierarchical figures in ASD have also provided conflicting results, although these studies have used selective and divided attention methodologies. It was predicted that, in an undirected task, the relative salience of the local elements over the global form would be greater in the ASD group compared to the control group. Participants in the control group were also predicted to be more sensitive to density changes than the ASD group, with an increasing bias towards the global form when a greater number of local elements are present.

5.5.2 Method

Materials

Standard and comparison stimuli consisted of global letters composed of different local letters (see Figure 5.15). A total of 12 stimuli types were created using the letters T, A, N, H, and F at both local and global levels (i.e., Ah, An, Fh, Ft, Ha, Hf, Hn, Ht, Na, Nh, Tf, Th). The letters were selected as they were relatively well matched on visual complexity and frequency (percent occurrence in the English language ranged from 2.3 to 9.3; Solso & King, 1976). The stimuli varied in density of local elements and were composed of four, five or six local elements in width and height, resulting in 36 stimuli in total (3 by 12). While the number of local elements changed, the size of the local elements and global form remained constant.



(a) (b) (c)

Figure 5.15 Example of stimuli used in the Navon Similarity Judgment Task. The Tf stimulus is presented at density 4 (a), 5 (b) and 6 (c).

Each global form measured approximately 4.5 cm in width and 5.5 cm in height, while each local element was 0.4 cm in width and 0.5 cm in height. At a typical viewing distance of 50 cm, the stimuli subtended approximately 6.3° of visual angle, and the local elements subtended 0.6°. This overall visual angle is within the limit in which a global advantage is typically detected in normal populations (Kinchla & Wolfe, 1979). As a preliminary check, the choice of stimulus size and the three density levels was piloted on 10 adult volunteers. Using the similarity judgment paradigm, nine participants showed a consistent preference to matching at the global level (selecting 31 or more global matches from 36), while one participant show a local preference (selecting one global match from 36). As a global bias was dominant in the adult sample, no changes were considered necessary to the stimuli.

Procedure

Participants were seated in front of the 15-inch desktop touch screen on which stimuli were presented. They were told to look at a red cross, positioned centrally in the top half of the screen, and were instructed as follows: *'Watch the cross carefully because after it disappears another figure will appear very quickly. Then two more figures will appear below. What I want you to do is pick one of the two bottom figures that you think was most like the top one. There are no right or wrong answers, as either one could be right. Don't take too long to think about it, just quickly choose the one you think was most like the top one."*

On each trial the red cross appeared on the screen for one second, followed by a blank screen for 150 ms, and was replaced by the standard figure which appeared for 250 ms (see Figure 5.16). The two comparison stimuli were presented in the bottom half of the screen and remained until the participant had made a response or four seconds had elapsed. One of the comparison figures had the same global letter as the standard but was composed of different local letters; the other comparison figure had the same local letters as the standard but they were arranged into a different global letter. Participants indicated their choice by touching one of the comparison figures.

If the participant did not make a response within the four second time limit an error sound was played (Windows sound file: utopia error.wav) and the following message appeared on the screen for two seconds: "*You took too long, please try again!*" The trial was repeated from the presentation of the red cross. If the participant touched an area of the screen that did not include a comparison figure the error sound was played with the message: "*Please touch one of the figures*". The message appeared for two seconds and the trial was resumed.

The task began with two practice trials followed by 36 test trials (the 12 stimuli presented at three levels of density). Within a single trial, the standard and comparison figures were of the same density. The stimuli triads were presented in a fixed random order for each participant. Each type of comparison figure (i.e., global match and local match to the standard) appeared equally often in the left and right positions. The number of global matches made by each participant was taken as a measure of processing bias, over the whole task and at each level of density.

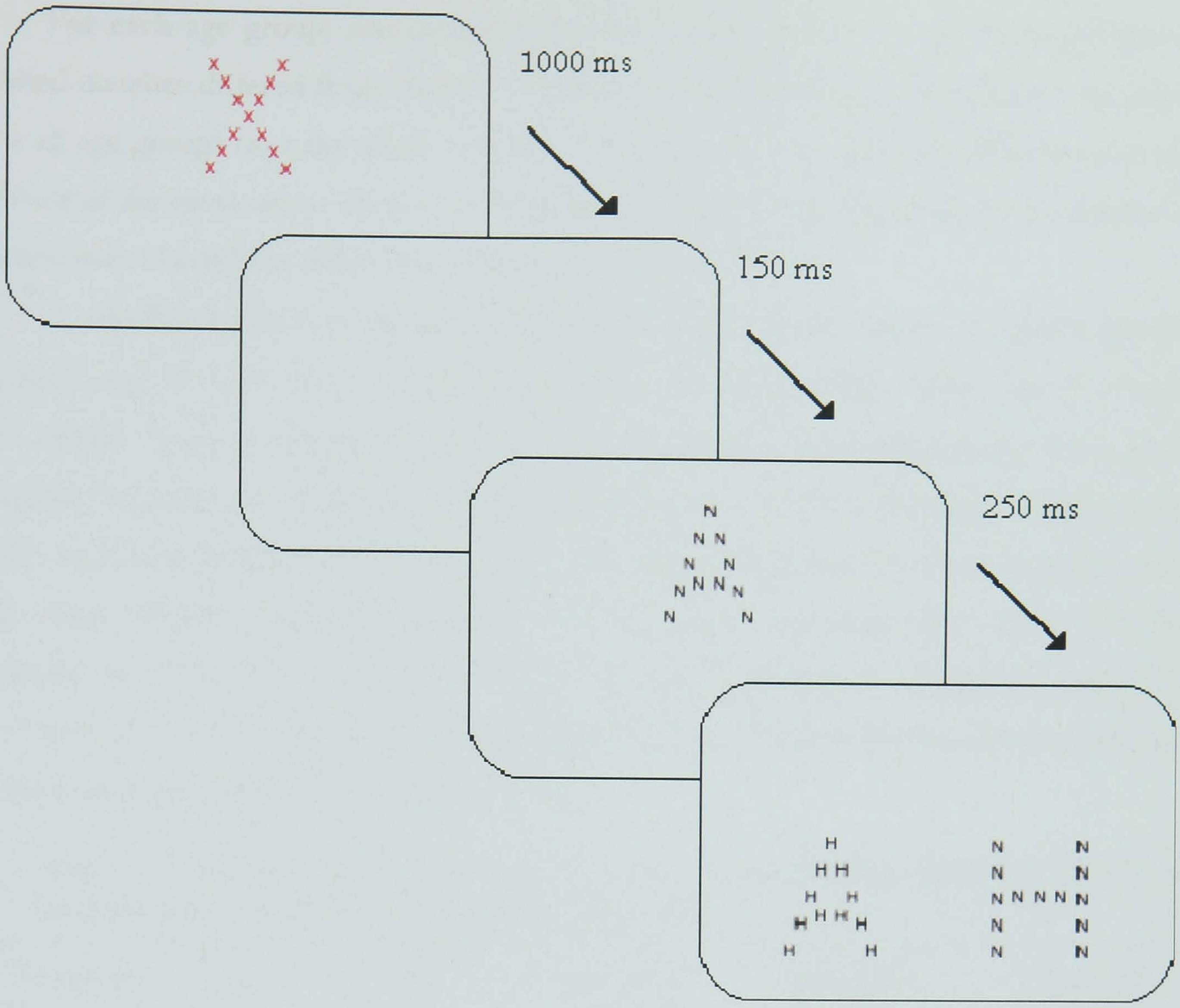


Figure 5.16 Sequential presentation of stimuli in the Navon Similarity Judgment Task, showing the orientation cross stimulus, an example of a standard stimulus and corresponding comparison stimuli.

5.5.3 Typical Development Results

All TD participants were administered the Navon Similarity Judgment Task (N = 204). A data file was lost for one female participant from the 11-13 year group; although her overall score was recovered, information at each level of stimulus density was absent. Twenty participants consistently selected a global match on all the 36 trials. These participants were mostly older: 11 from the 17-25 year group (6 male, 5 female), 6 from the 14-16 year group (2 male, 4 female) and three (male) participants from the 11-13 year group. At the other extreme, 4 participants consistently selected a comparison figure on all trials that had the same local letter as the standard. This group consisted of 3 participants from the 8-10 year group (1 male, 2 female) and one male from the 11-13 year group.

Table 5.15 shows the mean number of global matches made over the task and for each of the three density levels within each age group. The mean number of global matches over the whole task was negatively skewed for each age group (z-scores -2.2 to -8.8) showing the tendency towards making more global matches than local matches across all participants. Due to the non-normality of the data non-parametric statistical tests were used.

For each age group, one-sample t-tests were performed to verify if the number of global matches differed from chance. A significant bias toward global matches was found for all age groups over the whole task ($p < .01$, two-tailed). A lack of global bias was only shown at the most sparse density level for the 8-10 year group where the mean number of global matches did not differ from chance ($t_{(53)} = 1.60, p = .12$).

A significant effect of age group was found for the total number of global matches overall and at each level of stimulus density (Kruskal Wallis Tests, all $\chi^2 > 31.00$, $p < .0005$). In all conditions the two youngest age groups made significantly fewer global matches than the two oldest age groups (all $z > 2.70, p < .007$), indicating a developmental shift occurring between 11-13 years and 14-16 years. Male and female participants made the same number of global selections over the entire task (males $M = 27.8$, $SD = 9.6$; females $M = 26.4$, $SD = 10.9$; $z = 0.70$, $p = .49$) and at each level of stimulus density (all $z < 0.95$, $p > .34$). Male and female participants also did not differ in response selection within each age group (all $z < 1.64, p > .10$).

Table 5.15 Number of global matches made on the Navon Similarity Judgment Task overall and at the three density levels, by age group: Mean (*SD*).

Age group (years)	N	Total[a] (max = 36)	Density 4[b] (max = 12)	Density 5[c] (max = 12)	Density 6[d] (max = 12)
8-10	54	21.9 (10.8)	6.8 (3.6)	7.5 (4.0)	7.6 (3.7)
11-13	43†	24.4 (11.0)	7.2 (3.8)	8.5 (3.9)	8.4 (3.7)
14-16	51	30.1 (8.2)	9.4 (2.8)	10.4 (2.9)	10.3 (2.8)
17-25	56	31.2 (8.2)	10.0 (2.9)	10.7 (2.9)	10.5 (2.7)

†N = 42 at density 4, 5 and 6.
[a]8-10, 11-13 < 14-16, 17-25, $p < .003$
[b]8-10, 11-13 < 14-16, 17-25, $p < .003$
[c]8-10, 11-13 < 14-16, 17-25, $p < .004$
[d]8-10, 11-13 < 14-16, 17-25, $p < .007$

Significant effects of age group on the total number of global matches were found for each gender (males $\chi^2 = 18.99$, females $\chi^2 = 27.00$, $p < .0005$). Figure 5.17 shows this developmental effect for each level of stimulus density for male and females participants separately. As shown in the graph, males show a steady increase with age in the number of global matches, with significant increases between 8-10 years and 14-16 years (all $z > 2.32$, $p < .03$), and between 11-13 years and 17-25 years (all $z > 2.15$, $p < .04$). For female participants an increase in the number of global matches occurred between 11-13 years and

14-16 years (all $z > 2.39, p < .02$), while the two youngest groups and the two oldest groups did not differ (all $z > 0.99, p > .31$).

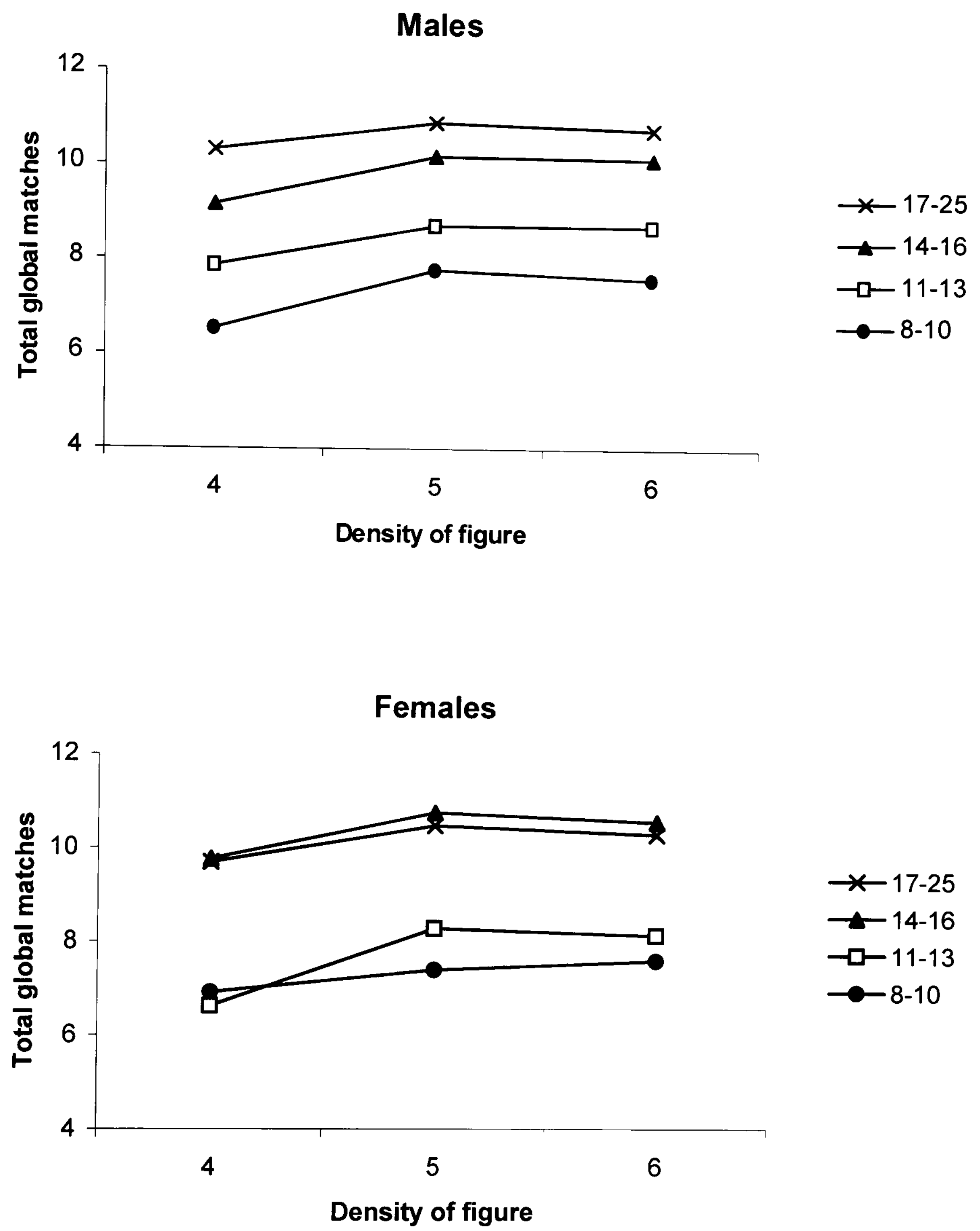


Figure 5.17 Mean number of global matches made at each level of stimulus density on the Navon Similarity Judgment Task by age group for males and females.

Effect of stimulus density

To assess the effect of stimulus density for each participant, Wilcoxon's Signed Ranks Tests compared the mean number of global matches made to the densest stimuli with the sparest stimuli. Over the whole sample significantly more global responses were made to density 6 compared to 4 ($z = 6.73$, N - ties = 147, $p < .0005$). This pattern was significant in each age group (all $z > 2.74$, $p < .006$). In terms of individual analyses, a high proportion of participants made more global matches at density level 6 than level 4: this was found for 61% (33 from 54) of the 8-10 year group, 64% (27 from 42) of the 11-13 year group and 59% (30 from 51) of the 14-16 year group. The percent was only 48% (27 from 56) of the 17-25 year olds, but a high proportion of participants (39%, 22 from 56) made an equal number of global matches at each density condition. A ceiling effect was therefore indicated in the oldest participants who tended to match by global form irrespective of the density of local elements.

Males and females were similarly sensitive to change in stimulus density, with significantly more global matches made to density 6 compared to 4 (males $z = 4.55$, N - ties = 68, $p < .0005$; females $z = 4.98$, N - ties = 79, $p < .0005$). As shown in Figure 5.17 this response pattern was apparent in each age group, with the exception of females in the 8-10 year group showing no difference in global response rate between the two densities ($z = 1.85$, N - ties = 28, $p = .07$). Males in the oldest age group also showed no difference in global matches between 4 and 6 ($z = 1.42$, N - ties = 15, $p = .16$).

Correlations

There was a significant positive correlation between age and the total number of global matches ($r_s = .46$, males $r_s = .48$; females $r_s = .43$; all $p < .0005$). A moderate positive correlation was found between IQ and preference towards global responding (FIQ $r_s = .20$, $p = .004$; FIQ-BD $r_s = .14$, $p = .04$; VIQ $r_s = .16$, $p = .02$; PIQ $r_s = .17$, $p = .02$). Interestingly this association was apparent in female participants (FIQ $r_s = .29$, $p = .002$; FIQ-BD $r_s = .21$, $p = .03$; VIQ $r_s = .25$, $p = .01$, PIQ $r_s = .22$, $p = .02$), but not in male participants (in whom $r_s = .06$ to $.11$, all $p > .27$). Fisher's z transformations did not find the correlation coefficients to differ between genders for each IQ measure (all $z_{r1-r2} < 1.40$, $p > .16$).

5.5.4 ASD and Control Results

The Navon Similarity Judgment Task was not administered to one 14-year-old in the ASD group, but data were available for all remaining participants in the ASD and control groups. The total number of global matches made over the task was negatively skewed for the control group (z-score = -4.3) and moderately less so for the ASD group (z-score = -1.7), showing the tendency for both groups to make relatively more global matches than local matches. As the data also failed tests of normality (Kolmogorov-Smirnov tests, all $p < .04$), non-parametric analyses were conducted.

Three participants from the control group only ever selected the global comparison figure, while no participants from the ASD group showed this consistent response pattern. No participants from either group showed consistent matching to the local element, although more ASD participants tended towards a local bias: a total of 7 from 30 ASD participants made 10 or fewer global matches, compared with 2 from 31 control participants (Fisher's Exact Test, p = .07, one-tailed).

The mean number of global matches over the task and at each level of stimulus density is shown in Table 5.16. One-sample t-tests were conducted to test whether the number of global matches made within each group differed from chance. Both groups showed a significant bias towards global matches over the task (controls $t_{(30)} = 6.36, p < .0005$; ASD $t_{(29)} = 2.04, p = .05$). A global bias was also found at all density levels for the control group (all $t_{(30)} > 5.10, p < .0005$), but only at density level 5 for the ASD group ($t_{(29)} = 2.5, p = .02$). The number of global matches failed to differ from chance at density 4 ($t_{(29)} = 1.57, p = .13$) and 6 ($t_{(29)} = 1.80, p = .08$) for the ASD group indicating a lack of global bias at the most sparse and dense stimulus conditions.

Table 5.16 Number of global matches made on the Navon Similarity Judgment Task, over the whole task and at three density levels, by group: Mean (*SD*).

	Group	N	Total (max = 36)	Density 4 (max = 12)	Density 5 (max = 12)	Density 6[a] (max = 12)
All	ASD	30	22.5 (12.1)	7.1 (3.9)	7.9 (4.2)	7.4 (4.3)
	Control	31	27.7 (8.5)	8.6 (2.9)	9.5 (3.0)	9.5 (3.1)
High IQ	ASD	24	23.5 (12.4)	7.4 (4.1)	8.4 (4.3)	7.8 (4.4)
	Control	25	28.7 (7.6)	8.9 (2.7)	9.9 (2.6)	9.9 (2.9)
Low IQ	ASD	6	18.5 (10.7)	6.2 (3.5)	6.2 (3.7)	6.2 (3.9)
	Control	6	23.5 (11.2)	7.7 (3.7)	7.8 (4.3)	8.0 (3.7)

[a] ASD < Control, p = .02; $ASD_{high} < Control_{high}$, p = .02

The ASD group made fewer global matches over the whole task than the control group, although statistically this difference did not reach significance ($z = 1.58$, $p = .06$, one-tailed). This pattern was also found at density level 4 ($z = 1.46$, $p = .07$) and 5 ($z = 1.17$, $p = .12$), and reached statistical significance at the most dense stimulus level, 6 ($z = 2.00$, $p = .02$). This finding is depicted in Figure 5.18. Although the mean age for both groups was 14.8 years (*SD* = 2.4), the ASD group performance is most similar to the 8-10 year age group from the TD sample, whereas the control group is comparable to the 14-16 year age group (cf. Figure 5.17).

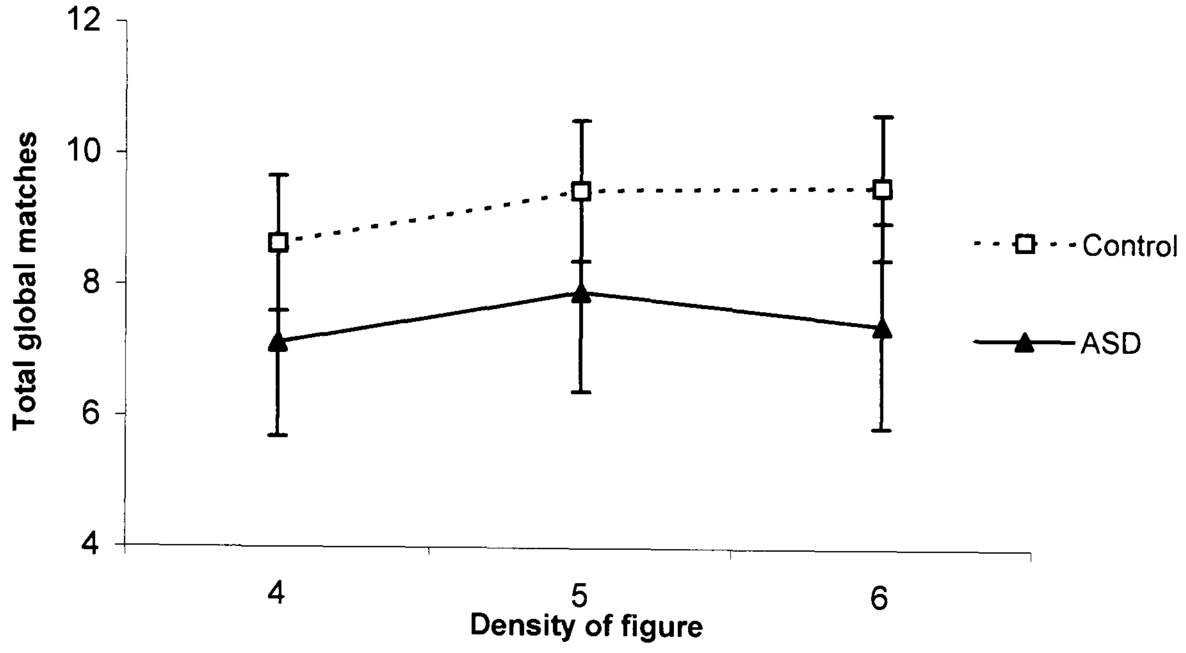


Figure 5.18 Mean number of global matches made at each level of stimulus density on the Navon Similarity Judgment Task, by group. Error bars show 95% confidence intervals.

No group differences were observed when comparing participants of high-ability on the total number of global matches made over the whole task ($z = 1.22, p = .11$, one-tailed) and at the density level 4 and 5 (all $z < 1.09, p > .14$). Significant differences were observed however at density level 6 ($z = 1.72$, $p = .04$). Participants of low-ability did not differ significantly on the number of global selections made overall ($z = 1.20, p = .12$), or at each level of stimulus density (all $z < 1.07, p > .14$).

Effect of stimulus density

Wilcoxon's Signed Rank Tests were used to assess sensitivity to stimulus density. The control group made significantly more global matches to the densest stimuli (level 6), compared to the sparsest (level 4; $z = 2.65$, N - ties = 23, $p = .008$, two-tailed). In contrast, the ASD group did not differ in the number of global matches made to the sparsest and the densest stimuli ($z = 1.09$, N - ties = 24, $p = .28$). On an individual basis, 61% of control participants (19 from 31) made more global matches at density level 6 compared to 4. This was apparent for half of the ASD participants (15 from 30; non-significant χ^2). Surprisingly the ASD group made significantly *fewer* global matches at the densest stimulus level 6 compared with level 5 ($z = 2.23$, N - ties = 20, $p = .03$).

As shown in Figure 5.19, the effect of density appears to be absent in individuals of low ability. In both low IQ groups, half of the participants (3 from 6) made more global matches at the denser stimulus level than the more sparse level. In the high IQ groups, 64% of control participants (16 from 25) made more global matches at density level 6 compared to 4, where as only 50% of ASD participants (12 from 24) showed this differentiation (non-significant χ^2).

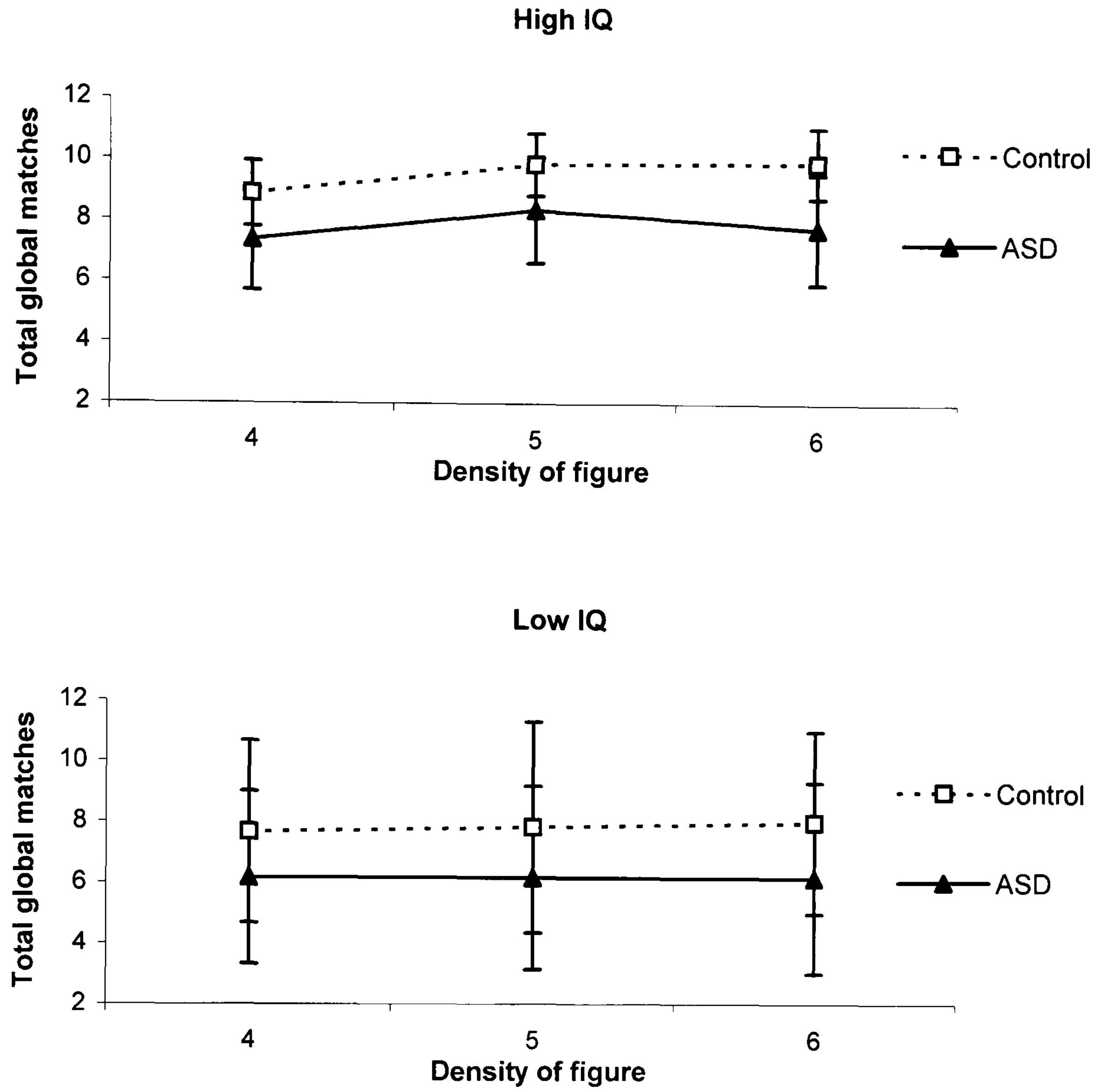


Figure 5.19 Mean number of global matches made at each level of stimulus density on the Navon Similarity Judgment Task, by groups divided by IQ. Error bars show 95% confidence intervals.

Correlations

No correlation was found between age and the number of global matches for the ASD group ($r_s = .01, p = .96$) or control group ($r_s = .07, p = .70$). A stronger positive correlation was observed between MA and total number of global matches, but this failed to reach statistical significance for both groups (ASD $r_s = .30, p = .10$; control $r_s = .17, p = .37$). IQ did not significantly relate to a global response bias in either group (control $r_s = .03$ to $.16$, all $p > .37$; ASD $r_s = .20$ to $.25$, all $p > .18$; all $z_{r1-r2} < 0.80, p > .42$).

5.5.5 *Summary of Navon Similarity Judgment Task*

As predicted TD participants were significantly biased towards making similarity judgments based on global form rather than local details. This global bias increased with age in both male and female participants, although an obvious developmental shift occurred in females between 11-13 and 14-16 years compared to the more continual increase in males. This difference may have been a result of the higher VIQ in females compared to males in the 14-16 year age group but unfortunately the nonparametric analyses precluded covarying IQ. IQ was also more strongly associated with global bias in females than males and this raises the question whether different mechanisms may operate in each gender. TD participants were sensitive to differences in stimulus density, making more global matches to stimuli of greater local element density. It was predicted that older participants would show greater sensitivity to density, and indeed, the differential between the most dense and sparse stimuli was not observed in females from the youngest age group. Participants in the oldest age group, however, were also not sensitive to changes in stimulus density. This finding appears to be a consequence of a strong global bias across all stimuli conditions in the oldest age group.

ASD participants were less biased towards making similarity judgments based on global form than control participants, especially at the most dense stimuli condition. They were not, however, significantly biased towards making similarity judgments based on local elements. This suggests that they may be less systematic in their response selection, not settling on a matching rule to adopt throughout the task. ASD participants were found to be less sensitive to changes in stimuli density than control participants.

The total number of global matches was selected as the main dependent variable from the Navon Similarity Judgment Task. This measure was taken as an indication of a processing bias towards strong central coherence. As significant differences were found between the ASD and controls groups at the most dense stimulus level, this variable was used to categorise participants as showing weak or strong coherence. Participants selecting 8 or fewer global matches (from 12) were classified as showing weaker central coherence compared to those who selected more than 8 global matches.

5.6 General Discussion

A summary of the findings from the four low-level visuo-spatial coherence tasks is presented in Table 5.17 for the TD sample and in Table 5.18 for the ASD versus control group comparisons (with effect sizes presented in Appendix Q). The percentage of participants in each group who were categorised as showing weak central coherence on each measure is shown in Table 5.19.

Performance on all four visuo-spatial measures showed developmental effects in the TD sample. Strong effects of age and IQ were found on tasks that demanded segmentation ability and local processing skills (EFT, Un/Segmented Block Design). A task requiring the opposite process of global coherence for rapid and accurate performance (Impossible Figures Task) was also strongly related to age and IQ. Perhaps surprisingly, developmental effects were also found on the Navon Similarity Judgment Task, a task designed to tap processing bias, rather than ability. On this measure, participants of higher developmental level showed greater bias towards matching by global form. As shown in Table 5.17, a substantial amount of residual variance remained after the effects of age, IQ and gender were removed, some of which may reflect individual differences in cognitive style.

An effect of gender was found on the Impossible Figures Task, where males surpassed females in the ability to differentiate between possible and impossible geometric figures. Contrary to predictions, gender effects were not found on the Un/segmented Block Design and the EFT, although the developmental trajectories appeared to differ for males and females. In particular, females were shown to improve in performance at an earlier age than males. An interesting question is whether this finding may relate to differences in the development of brain structures such as the frontal lobe and the corpus callosum, which may differ between genders (Giedd, Blumenthal, Jeffries, Rajapakse et al., 1999).

Predictions from weak central coherence theory were supported in three of the four low-level visuo-spatial measures of coherence. Compared to their age- and IQ-matched controls, the ASD group were less biased towards matching by global form on the Navon Similarity Judgment Task, showed superior disembedding skills and greater ability to ignore the global gestalt on the EFT, and were less able to determine geometric impossibility on the Impossible Figures Task.

A relative lack of benefit from segmentation on the Un/segmented Block Design Test was not confirmed in the ASD group, despite the consistent finding of superior performance on Block Design tasks in the autism literature. This lack of effect may stem from heterogeneity in the sample of individuals with ASD selected for this study. It is worth noting that the subgroup of individuals with ASD who had peak Block Design performance *did* show reduced benefit from segmentation. As a further means of analysis, the frequency of configuration-breaking errors was compared between the ASD and control groups, but no differences were found. However, the frequency of configuration-breaking errors was not considered in relation to all errors made by the individual. The occurrence of detail-violating errors (e.g., a rotated single block) was not collated in the present study and limits the conclusions that can be drawn. Comparing rates of detail-violating errors relative to configural-breaking errors will be considered in future studies.

The relationship between all four low-level visuo-spatial coherence measures presented in this chapter will be examined in Chapter 9. Cross-modality processing will also be investigated by comparing performance between low-level tasks in the auditory domain and high-level tasks in verbal and visuo-spatial domains.

Table 5.17 Summary of low-level visuo-spatial task findings in the TD sample.

Visuo-spatial Task	Dependent Variable (Direction of high values)	Age effects?	IQ effects?	Gender effects?	Proportion of variance not explained by age, IQ and gender	Interactions
Un/segmented Block Design	Relative benefit from segmentation (Strong CC)	Yes, Negative	Yes, Negative	No	79%	F > M at 8-10 (not after covary FIQ)
Embedded Figures Test	Detection time (Strong CC)	Yes, Negative	Yes, Negative	No	73%	F > M at 8-10 M > F at 11-13
Impossible Figures Task	A′ (Strong CC)	Yes, Positive	Yes, Positive	Yes, M > F	80%	M > F at 17-25 M > F at 8-10 (trend)
Navon Similarity Judgment Task	Total global matches (Strong CC)	Yes, Positive	Yes, Positive	No	87%	Positive correlation with IQ in F only

Note: M = males, F = females

Table 5.18 Summary of low-level visuo-spatial task findings in the ASD and control groups.

Visuo-spatial Task	Dependent Variable (Direction of high values)	Group effects?	Age effects?	MA effects?	IQ effects?	Other comments
Un/segmented Block Design	Relative benefit from segmentation (Strong CC)	No	Yes, positive in ASD only	No	Yes, Negative	No correlation with FIQ-BD
Embedded Figures Test	Detection time (Strong CC)	ASD < Control	No	Yes, negative in Controls only	Yes, negative in Controls only	$\text{ASD}_{\text{high}} < \text{Control}_{\text{high}}$ $\text{ASD}_{\text{low}} < \text{Control}_{\text{low}}$
Impossible Figures Task	A′ (Strong CC)	ASD < Control	No	Yes, positive	Yes, positive	
Navon Similarity Judgment Task	Total global matches (Strong CC)	ASD < Control (at density 6)	No	No	No	$\text{ASD}_{\text{high}} < \text{Control}_{\text{high}}$ (at density 6)

Table 5.19 Percentage (and N/total N) of participants in each group who showed weak central coherence on each task.

Visuo-spatial Task	Marker for weak CC	8-10	11-13	14-16	17-25	ASD	Control
Un/segmented Block Design[a]	Relative benefit of segmentation < 30%	4% (2/54)	16% (7/43)	24% (12/51)	27% (15/56)	23% (7/31)	13% (4/31)
Embedded Figures Test[b]	Detection time < 20 secs	15% (8/54)	23% (10/43)	43% (22/51)	61% (34/56)	36% (11/31)	20% (6/30)
Impossible Figures Task[c]	A' < .80	35% (19/54)	10% (4/42)	12% (6/50)	11% (6/56)	24% (7/29)	21% (6/28)
Navon Similarity Judgment Task[d]	Global matches at most dense stimuli < 8 (from 12)	54% (29/54)	36% (15/42)	14% (7/51)	13% (7/56)	43% (13/30)	26% (8/31)

[a]8-10 < 11-13, 14-16, 17-25, $p < .04$
[b]8-10 < 14-16, 17-25, $p < .008$; 11-13 < 17-25, $p < .0005$
[c]8-10 > 11-13, 17-25, $p < .02$
[d]8-10 > 14-16, 17-25, $p < .002$, 11-13 > 17-25, $p = .03$

Chapter 6. Low-level auditory coherence tasks

6.1 Introduction

This chapter presents experimental findings from four tasks designed to assess individual differences in the processing of low-level auditory material. Two tasks require the ability to isolate a sound from its embedded context and demand local processing for successful performance (Phoneme Segmentation, Chord Segmentation). The third task, in which a local bias is also advantageous, is the identification of the absolute properties of music stimuli (Pitch Identification). In order to determine whether exact perception is special to music, comparison is made with the ability to retain stable pairings in the visual modality (Line Orientation). The final task assesses preference for local versus global endings of musical sequences (Chord Sequence Task), which was designed to measure pure processing bias, rather than ability. Two independent measures of local and global processing are obtained from this latter task.

As three of the four auditory tasks involve the manipulation of musical stimuli, it was acknowledged that the formal musical training of an individual might influence task performance. Information was therefore gathered from participants on their degree of musical training and experience (see Appendix E).

6.2 Phoneme Segmentation Task

Phonological awareness has been defined as the sensitivity to sounds in spoken language (Gallagher, 1995). This awareness can be assessed at the level of the syllable (e.g., syllable tapping tasks), onset or rime (e.g., alliteration and rhyme tasks), or the phoneme (e.g., phoneme manipulation tasks such as deletion, insertion, substitution). The hardest level of analysis is reported to be at the phonemic level, which is known to develop in conjunction with reading ability (Goswami & Bryant, 1990). Children with specific language impairments (SLI) are claimed to persist much longer than other children in analysing speech at the level of syllabic units, without awareness of phonemes (Bishop, 1997).

A phoneme is regarded as the smallest unit of speech that serves to distinguish one word from another in a language. Early models of speech perception describe a "bottom-up" process involved in transforming a sound wave into meaning. A series of stages are described in which acoustic signals are extracted and stored in sensory memory and then mapped onto linguistic information in the form of phonemes (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). This model, however, has been regarded by many psycholinguistics as too simplistic (Bishop, 1997). It would assume that children must first learn to recognise phonemes when acquiring new words. However, young children can successfully recognise and produce words without necessarily being able to match or identify individual phonemes. The awareness of phonemic elements may not occur until they are exposed to printed words (Bishop, 1997).

Several lines of reasoning suggest that the phoneme is not a natural or obvious unit of speech perception and can require active segmentation of linguistic information for identification. First, phonemes are not discrete entities, but various phonemes are said to blend into one another. The actual sound produced for one phoneme is determined by the context of the surrounding phonemes. This phenomenon, known as *co-articulation*, results in the "smearing together of information in the acoustic signal about individual phonemes" (Nygaard & Pisoni, 1995, p. 66). A classic study conducted by Savin and Bever (1970) also provides evidence to suggest the phoneme is an "abstract" linguistic unit, rather than a perceptual entity. Undergraduate students were required to listen for initial phonemes (e.g., */p/*) or syllables (e.g., *ba*) in lists of syllables. Their main finding was that participants responded faster when identifying syllables than when identifying phonemes within syllables. Savin and Bever concluded, "phonemes are perceived only by an analysis of already perceived syllables" (p. 300). These studies suggest that the identification of a single phoneme within a speech stream is a difficult task.

The finding that a smaller unit (the phoneme) is identified by fractionating or segmenting the larger unit (the syllable or word) suggests a plausible verbal analogy to local and global processing in the visual domain. Thus, identifying a phoneme within a word may be akin to spotting a hidden figure in the EFT, for example. A measure of auditory segmentation was therefore devised for the present study that required the ability to identify a target phoneme within a non-word. Although real words may provide a stronger gestalt (comparable to meaningful contexts in visual processing tasks such as the CEFT), non-words were selected as the whole form to ensure that word familiarity was consistent across participants and to minimise the effect of individual differences in literacy/spelling ability between participants.

Phoneme segmentation in ASD

Phonological processing has been examined to some extent in individuals with ASD within the study of language impairment. To date, there has been no suggestion of superior segmentation of linguistic material in this clinical group. Early investigations into the acquisition of phonology, syntax, and morphology, demonstrated that children with autism did not differ from age-, IQ-, or language-matched comparison groups (Bartolucci, Pierce, Streiner, & Eppel, 1976; Bartolucci & Pierce, 1977; Pierce & Bartolucci, 1977). It is well known that children with autism vary greatly in the onset of language and rate of development, and it has been suggested that this heterogeneity was not fully addressed in these early studies (Roberts, Rice, & Tager-Flusberg, 2004). Furthermore, phonological awareness was typically assessed through the production and perception of speech sounds (such as the ability to discriminate between words differing by one phoneme; e.g., /bowt/ /dowt/; Bartolucci & Pierce, 1977), and not phoneme segmentation skill per se.

In a recent large-scale study of language processing in children with autism (N = 89, 4 to 14 years), Kjelgaard and Tager-Flusberg (2001) reported significant variation in language abilities, with approximately 25% of the sample having normal language skills. The language-impaired children with autism showed a similar profile of performance across language measures to children with specific language impairment (SLI). In particular, these children were impaired on a non-word repetition task, a measure that does not rely solely on rote memory, but on the accurate perception and encoding of phonological information (Bishop, 1997).

To date, one study has explicitly studied phoneme segmentation in autism. Scheuffgen (1998) demonstrated that a group of children with ASD (N = 11, mean age 11 years) were comparable to a younger reading-level matched control group (mean age 8 years) on measures of rhyming and non-word reading, but were not as capable as the controls in performing a spoonerism task. It was suggested that the working memory requirement of the spoonerism task (i.e., to hold in mind and switch the initial phonemes of a word pair) over and above the segmentation skill involved, accounted for the group differences and that the general phonological skills of the ASD group were intact.

Predictions

In the Phoneme Segmentation Task, developed for the present study, participants were required to detect the presence of a target phoneme within a non-word. Two conditions of the task were contrasted: (a) trials when the phoneme was presented as the initial sound, and (b) trials when the phoneme was "embedded" as the medial or final sound. It was predicted that weak central coherence would be evidenced by an enhanced ability to segment linguistic information, resulting from a bias towards processing features rather than the whole form. Individuals with weak coherence are expected to show less effect of the embedding context, reflected in better performance overall, and less effect of target position. In contrast, a relative difficulty of isolating a phoneme positioned as the medial or final sound, as compared to the initial sound, is suggested to reflect strong coherence.

Phonological awareness has been shown to emerge developmentally along a continuum, from early infancy through to approximately age seven years (Norris & Hoffman, 2002). As the youngest TD participants in the present study are older than the typical age of development in phonological awareness, no specific effects of development are expected. Individual differences in phoneme segmentation are, however, predicted, in accordance with individual differences in coherence processing style.

The existing literature on phonological processing in ASD may suggest that individuals with ASD and language impairment would show deficits in the perception and encoding of phonological information. Although weak central coherence theory would predict good disembedding performance regardless of verbal abilities, superior performance on the Phoneme Segmentation Task may be shown only in a subgroup of individuals with ASD that have preserved language skills.

6.2.2 Method

Materials

Forty-five non-words were used, 15 in each of three conditions: (1) target as initial sound, (2) target as medial or final sound, and (3) target absent. An equal number of one-, two-, and three-syllable non-words were present in each condition. For simplicity one target phoneme /p/ was used throughout the task. The phoneme was either placed next to a vowel (e.g., /ip/ /pa/), or as part of a consonant cluster (e.g., /plo/, /asp/), with each type equally represented in the initial and medial/final position conditions. This design was to ensure a range of difficulty as previous studies have confirmed that phoneme detection rates depend on the complexity of the syllable in which the target appears (Cutler, Mehler, Norris, & Segui, 1987). Furthermore, plausible foils included other consonant stops (e.g., /b/, /t/, /d/) to increase the difficulty of the task.

The stimuli were pre-recorded to ensure consistency of presentation and to eliminate visual cues, which can occur if presented by live voice. A female speaker of British English initially read the list of non-words in a quiet room. Her speech was recorded using Sony Minidisk recorder and each word was stored as an individual sound file. The presentation of stimuli was controlled by SuperLab Pro software run by a laptop computer and played through high quality speakers (Harman Kardon). The response input was a two-button serial mouse that had "No" and "Yes" written on the left and right buttons respectively.

The test items were selected from 63 non-words created by the author that had been piloted on a sample of ten adults. Accuracy rates in detecting the target phoneme were collected with the aim to select items that demonstrated a range of difficulty in order to avoid ceiling and floor effects. The final stimuli set and the pilot data for these items is presented in Appendix H.

Procedure

Participants were seated in front of the desktop touch screen and two speakers that were attached to the laptop computer. The experimenter explained to the participant that they would soon hear some words that were 'made up' and had no meaning. Their task was to listen for a /p/ sound that may occur in the word. The experimenter orally presented examples of non-words that contained the target phoneme at each possible location (initial *plo*, final *vip*, medial *lipod*) as well as the target being absent (*dowen*). Participants were instructed to press the right button (Yes) if they heard the /p/ sound within the non-word or press the left button (No) if they did not hear it.

Three practice items with corrective feedback were then presented (*pollan, yeat, noop*). Each trial began with the word "next" appearing on the touch screen. The participant controlled the presentation of each item by pressing either mouse button when they were ready to hear the next word. If an error was made on a practice item, the words "*Listen again*" appeared and the trial was repeated.

Participants had the option to repeat the practice phase (with the same items) before starting the test items. The main task was presented in three blocks of 15 trials, with the test items presented in a fixed random order. Each block was separated by a reminder of the task requirements: "*Listen for this sound in the word*" appeared on the screen, followed by the /p/ sound. Participants had the option to hear the /p/ sound again, or to continue with the task. No feedback was given in the experimental phase and participants had unlimited time to make their response. Reaction time was measured from the offset of each non-word.

6.2.3 *Typical Development Results*

All TD participants were administered the Phoneme Segmentation Task. Correct and incorrect detections of the target phoneme were subjected to signal detection analysis to examine participants' ability to discriminate target absent from target present trials and to measure response bias. All values for A', the non-parametric index of sensitivity (Grier, 1971), were greater than 0.50 (range 0.67 to 0.98), suggesting that no individual was performing at chance or had inadvertently mixed the Yes and No response buttons when responding. Individual data files from participants with A' values classed as outliers (i.e., values 3 *SD* below age group mean) were checked for any systematic response patterns that may indicate the participant was not performing the intended task (e.g., simply alternating between left and right keys). No invalid response patterns were detected and so no data were excluded. Examination of results showed no change when participants with ESL were excluded, so data from all participants were analysed.

The distribution of accuracy rates and A' were negatively skewed for each age group, indicating the tendency towards making correct judgments. All accuracy data failed Kolmogorov-Smirnov tests of normality and non-parametric statistical tests were therefore used and all outliers remained in the analysis.

As no time limit was imposed, it was necessary to examine the distribution of response times from all participants. A small cluster of response times (0.5%) were observed to be above 10 seconds and it was considered appropriate to cap these extreme outliers to 10 seconds. A logarithmic transformation of the time data was performed to normalise the distribution of scores and allow parametric statistical analyses.

Accuracy

Age groups did not differ significantly on the non-parametric measure of bias, B" ($\chi^2 = 4.29$, $p = .23$). Although 67% of test items contained the target phoneme, no bias towards responding to either the right or left key was apparent over the whole sample.

The number of correct responses in each condition for each age group is presented in Table 6.1 and Figure 6.1. Wilcoxon's Signed Ranks Tests revealed that participants in all age groups made more correct judgments when the target phoneme was the initial sound compared to when presented as the medial or final sound (all $z > 5.49$, $p < .0005$). This pattern was confirmed for male and female participants overall, and within each age group (all $z > 3.14$, $p < .003$).

No significant age group effects were found in each of the three conditions ($\chi^2 < 5.45$, $p > .14$), but were detected on the non-parametric measure of sensitivity A' ($\chi^2 = 7.82$, $p = .05$). The youngest age group was less able to discriminate between target present and target absent trials than the oldest age group ($z = 2.72$, $p = .007$). The oldest age group was also marginally better at the phoneme segmentation task than the 14-16 year olds ($z = 1.87$, $p = .06$). Significant effects of age group were not detected when analysing each gender separately (males $\chi^2 = 3.55$, $p = .32$; females $\chi^2 = 6.74$, $p = .08$).

There was no effect of gender on accuracy, with male and female participants making the same number of correct detections in each condition (all $z < 1.10$, $p > .26$). Furthermore, no differences were found between male and female participants overall and within each age group on the sensitivity measure A' (all $z < 2.3$, $p > .24$).

Table 6.1 Number of correct responses for each condition of the Phoneme Segmentation Task (maximum = 15) and the overall sensitivity index, A' for each age group: Mean (*SD*).

Age group (years)	N	Target absent	Target initial	Target medial/final	A'[a]
8-10	54	11.9 (1.9)	14.0 (1.4)	10.9 (2.4)	0.89 (0.04)
11-13	43	12.3 (1.6)	14.0 (1.2)	10.7 (1.8)	0.89 (0.04)
14-16	51	12.5 (1.6)	13.9 (1.4)	10.2 (2.3)	0.89 (0.05)
17-25	56	12.5 (1.3)	14.3 (1.4)	10.8 (2.7)	0.90 (0.06)

[a] 8-10 < 17-25, p = .007

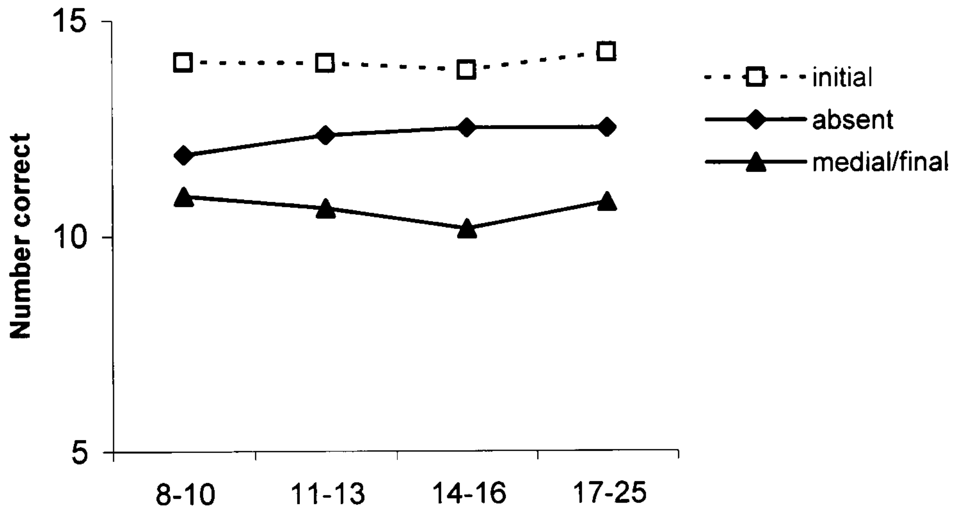


Figure 6.1 Mean number of correct responses for each condition of the Phoneme Segmentation Task (target absent, target initial sound, target medial/final sound; max = 15) for each age group.

Response time

Table 6.2 presents the mean response time for correct detections of the target phoneme in each of the three experimental conditions for each age group. As mean response time, irrespective of correctness, yielded the same pattern of results as mean response time for correct detections, the latter variable was used for clarity. Statistical analyses were performed on the logarithmic transformed time data.

A two-way ANOVA was conducted on the response time for target absent trials. A significant effect of age group was found ($F_{(3,196)}$ = 14.50, p < .0005), with the youngest group taking significantly longer to judge the absence of the phoneme than all older age groups (Tukey's HSD, all p < .007). No main effect of gender was found ($F_{(1,196)}$ = 0.74, p = .39), nor interaction between age group and gender ($F_{(3,196)}$ = 0.59, p = .62).

Table 6.2 Response time (in seconds) for correct detections in each condition of the Phoneme Segmentation Task for each age group: Mean (*SD*).

Age group (years)	N	Target absent[a]	Target initial[b]	Target medial/final[c]
8-10	54	1.76 (1.07)	1.13 (0.64)	1.38 (0.63)
11-13	43	1.32 (1.06)	0.86 (0.47)	1.13 (0.57)
14-16	51	1.01 (0.39)	0.88 (0.48)	1.10 (0.64)
17-25	56	1.07 (0.45)	0.85 (0.52)	1.02 (0.58)

[a]8-10 > 11-13, 14-16, 17-25, $p < .007$
[b]8-10 > 17-25, $p = .03$
[c]8-10 > 17-25, $p = .01$

As the effect of phoneme position (i.e., degree of embedding) was of interest to the present study, correct response times for target present trials was analysed with a two-way repeated measures ANOVA with age group and gender as between-subjects factors, and phoneme position (initial vs. medial/final) as the within-subject factor. There was a main effect of age group ($F_{(3,196)} = 4.07$, $p = .008$), with the youngest group taking longer to respond overall compared to the two oldest groups (Tukey's HSD, all $p < .03$). There was no significant main effect of gender ($F_{(1,196)} = 0.55$, $p = .46$), and no interaction between gender and age group ($F_{(3,196)} = 0.95$, $p = .42$). A main effect of phoneme position was detected ($F_{(1,196)} = 69.37$, $p < .0005$), with longer response times for the medial/final positions of the target phoneme than initial positions. There was no interaction between phoneme position and age group ($F_{(3,196)} = 0.37$, $p = .78$), phoneme position and gender ($F_{(1,196)} = 1.05$, $p = .31$), or between phoneme position, age group and gender ($F_{(3,196)} = 0.19$, $p = .90$). The main effect of phoneme position was therefore consistent across age groups, and across male and female participants within each age group.

Relative effect of phoneme position

To obtain an index of the relative difficulty of disembedding a target phoneme, a score was calculated based on the mean difference in correct response times between phoneme conditions (medial/final minus initial), divided by the mean time for medial/final phoneme. As the index was negatively skewed for each age group (z-scores -3.7 to -8.0), non-parametric analyses were used. No effect of age group was evident on the phoneme position index ($\chi^2 = 1.64$, $p = .65$), in either male ($\chi^2 = 2.00$, $p = .57$) or female participants ($\chi^2 = 1.19$, $p = .76$). No gender effects were apparent overall ($z = 0.75$, $p = .45$), or in each age group (all $z < 0.95$, $p > .33$). As shown in Figure 6.2 the effect of phoneme position appears to be greater for males than females, with a stronger effect in the older age groups, however the large variance in the data occludes any group differences.

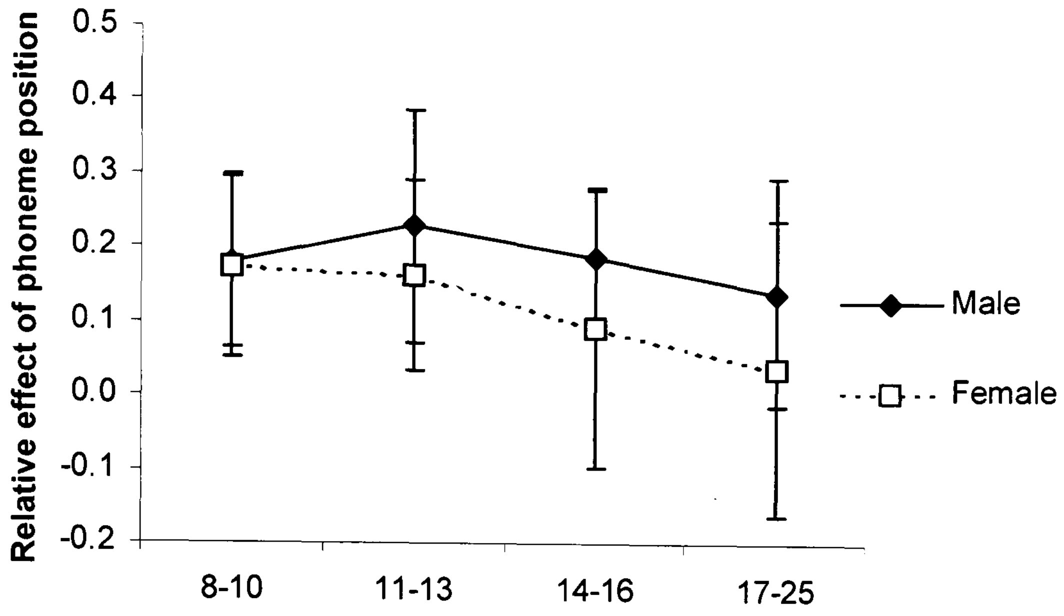


Figure 6.2 Relative effect of phoneme position on response time in the Phoneme Segmentation Task for male and female participants in each age group. Error bars show 95% confidence intervals

Correlations

There was a significant positive correlation between A' and age (r_s = .18, p = .009). This relationship was significant in female participants (r_s = .24, p = .01), but not male participants (r_s = .11, p = .30). Fisher's z transformation did not confirm that the coefficients were significantly different however (z_{r1-r2} = 0.94, p = .35). A positive correlation was found between A' and all IQ indices (FIQ r_s = .29, FIQ-BD r_s = .27, VIQ r_s = .26, PIQ r_s = .23, all $p < .001$). This association between task performance and IQ was apparent in male participants (r_s = .28 to .33, $p < .006$) and also in female participants (r_s = .21 to .25, $p < .03$, with the exception of PIQ r_s = .16, p = .10).

The relative effect of position on response time ($RT_{medial/final} - RT_{intial} / RT_{medial/final}$) did not correlate with age in the TD sample (r_s = -.06, p = .39, males r = -.01, p = .94, females r_s = -.13, p = .20). Furthermore, the relative effect of position index did not correlate with any IQ measure over the entire sample (r_s = -.05 to .04, $p > .51$). A trend was apparent in females for a negative correlation between the relative effect of position and FIQ (r_s = -.19, p = .06), indicating that females of high IQ showed less effect of phoneme position. Fisher's z transformation confirmed that the strength of the correlation coefficient for females differed significantly from that for males (r_s = .13, p = .22; z_{r1-r2} = 2.27, p = .02).

6.2.4 ASD and Control Results

All ASD and control participants were administered the Phoneme Segmentation Task. Two outliers were determined on the sensitivity measure, with A' values of .60 (control participant) and .53 (ASD participant), revealing a performance level that was close to chance. On examination of their individual data files it was clear that both participants were not performing the task (i.e., the ASD participant systematically alternated between left and right buttons, the control participant made straight runs of 10 left button presses and then 10 right button presses). The data from these two participants were therefore removed from the analyses. As both participants were of low IQ (FIQ 53 and 51) and of similar age (14.5 years and 13.7 years), their exclusion did not effect the age and IQ matching of the two groups.

Accuracy

The ASD and control groups did not demonstrate a bias towards responding to the left or right key overall, with no group differences on non-parametric measure of bias, B'' ($z = 0.19$, $p = .85$). The mean number of correct judgments for each of the three conditions (target absent, target initial, target medial/final) and the signal detection parameter of sensitivity (A') for each group is presented in Table 6.3. Groups did not differ on the number of correct judgments in each condition (all $z < 1.16$, $p > .12$, one-tailed). A significant difference was found between groups on A' ($z = 1.63$, $p = .05$), indicating that the ASD group were less able to discriminate between target absent and target present trials than the control group. This result was also found when participants of high IQ only were considered ($z = 2.01$, $p = .02$) and is opposite to the predictions of weak central coherence with ASD participants showing poorer ability to determine the presence of the phoneme than control participants.

Although the number of participants in each group was low (N = 5), low IQ ASD participants were able to correctly identify the absence of the phoneme more often than their matched controls ($z = 1.91$, $p = .03$). This finding may relate, however, to a difference in response bias between the two groups, which approached statistical significance ($z = 1.89$, $p = .06$). The low IQ ASD group had a positive bias score ($M = 0.38$, $SD = 0.58$) indicating a tendency towards stating the absence of the phoneme, while the low IQ control group has a negative bias score ($M = -0.30$. $SD = 0.32$). No significant group differences were found within participants with low IQ in detecting the presence of the phoneme in the initial ($z = 1.35$, $p = .09$) or medial/final positions ($z = 1.38$, $p = .08$), or in the overall measure of discrimination, A' ($z = 0.74$, $p = .23$).

Participants in each group correctly detected the phoneme more often when it was in the initial position than when it was in the medial or final position (Wilcoxon's Signed Ranks Test, all $z > 4.24$, $p < .0005$). Only one participant from the ASD group made fewer correct judgments when the phoneme was in the initial position than the medial/final, and no participants from the control group showed this pattern.

Table 6.3 Number of correct responses for each condition of the Phoneme Segmentation Task (maximum = 15) and the overall sensitivity index, A' for each group: Mean (*SD*).

	Group	N	Target absent[a]	Target initial	Target medial/final	A'[b]
All	ASD	30	11.7 (1.8)	13.2 (1.6)	9.4 (2.7)	0.84 (0.07)
	Control	30	12.1 (1.9)	13.6 (1.4)	9.7 (3.0)	0.87 (0.07)
High IQ	ASD	25	11.6 (1.9)	13.3 (1.8)	10.0 (1.9)	0.85 (0.07)
	Control	25	12.7 (1.3)	13.6 (1.4)	9.8 (3.2)	0.89 (0.05)
Low IQ	ASD	5	12.2 (1.9)	12.8 (0.8)	6.2 (4.1)	0.82 (0.06)
	Control	5	9.4 (1.9)	13.6 (0.9)	9.4 (1.9)	0.78 (0.07)

[a] $ASD_{high} < Control_{high}$, $p = .01$; $ASD_{low} > Control_{low}$, $p = .03$
[b] ASD < Control, $p = .05$; $ASD_{high} < Control_{high}$, $p = .02$

Effect of language development

As poor phonological awareness is often associated with language delay, it was predicted that a subgroup of ASD participants without early language impairment might show superior segmentation skills on the Phoneme Segmentation Task. Six individuals with ASD were identified as having delayed language development (with mean FIQ 73, VIQ 74, PIQ 78). When these six individuals, along with their matched controls, were removed from analysis no group differences were found on A' (ASD $M = 0.86$, $SD = 0.06$; control $M = 0.87$, $SD = 0.07$; $z = 1.29$, $p = .10$, one-tailed), or on the three conditions of phoneme position (all $z < 1.33$, $p > .09$).

Response time

Table 6.4 presents the mean response time for correct detections of the target phoneme in each of the three conditions for each group. No group differences were found on the time to correctly detect the absence of a phoneme ($t_{(58)} = 0.56$, $p = .58$; high IQ $t_{(48)} = 1.31$, $p = .20$; low IQ $z = 0.73$, $p = .47$).

To study the effect of phoneme position, correct response times (logarithmic transformed) were analysed with a repeated measures ANOVA with group as the between-subjects factor and phoneme position (initial vs. medial/final) as the within-subject factor. There was no main effect of group ($F_{(1,58)} = 0.05$, $p = .82$). A main effect of phoneme position was found ($F_{(1,58)} = 10.26$, $p = .002$) with participants taking longer to detect the phoneme when presented in medial/final positions than the initial position. A significant interaction between phoneme position and group was found ($F_{(1,58)} = 5.44$, $p = .01$, one-tailed). As depicted in Figure 6.3, the effect of phoneme position was not as pronounced in the ASD group as compared to the control group. The processing time to determine the presence of the target phoneme when presented as the initial sound compared to when it was the medial/final sound appears to be equivalent in the ASD group. This was confirmed by paired samples t-tests showing a significant difference between conditions for the control group ($t_{(29)} = 3.75$, $p = .001$) and no difference for the ASD group ($t_{(29)} = 0.14$, $p = .89$). The same pattern of results was found when excluding participants of low IQ ($F_{(1,48)} = 3.80$, $p = .03$), or with early language delay ($F_{(1,46)} = 2.81$, $p = .05$).

Table 6.4 Response time (in seconds) for correct detections in each condition of the Phoneme Segmentation Task for each group: Mean (*SD*).

	Group	N	Target absent	Target initial	Target medial/final
All	ASD	30	1.52 (1.02)	1.39 (0.94)	1.51 (1.14)
	Control	30	1.40 (0.90)	1.20 (0.87)	1.60 (1.13)
High IQ	ASD	25	1.45 (0.74)	1.41 (0.93)	1.46 (0.89)
	Control	25	1.20 (0.66)	1.05 (0.64)	1.34 (0.76)
Low IQ	ASD	5	1.90 (2.01)	1.29 (1.10)	1.74 (2.15)
	Control	5	2.40 (1.33)	1.94 (1.49)	2.93 (1.79)

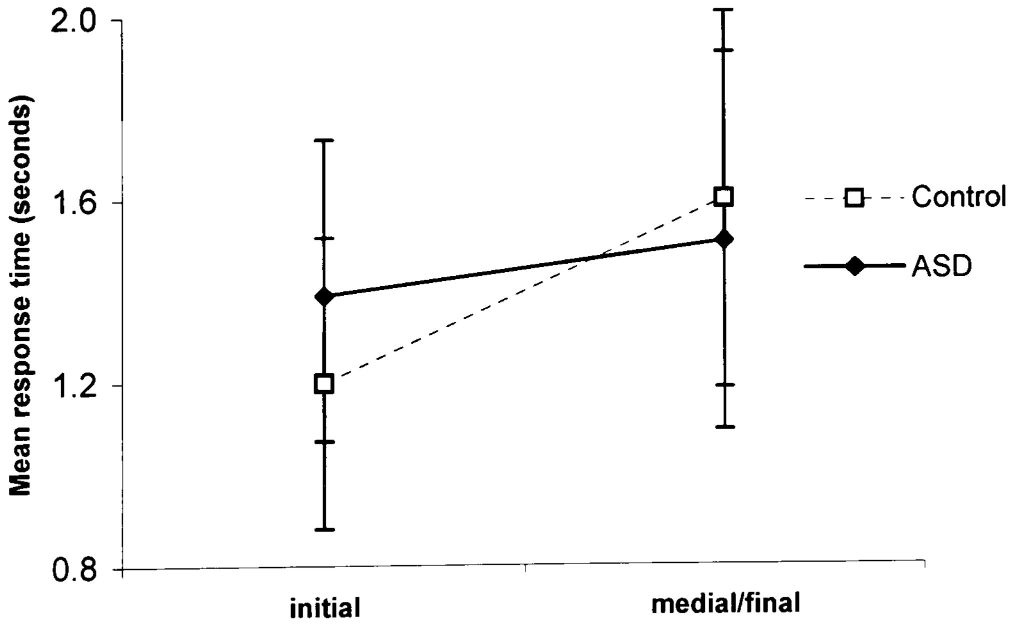


Figure 6.3 Mean response time (in seconds) for correct detection of the phoneme at initial and medial/final positions for each group. Error bars show 95% confidence intervals.

Relative effect of phoneme position

The relative difficulty in identifying an embedded phoneme in the medial/final position compared to the initial position was calculated for each participant ($RT_{medial\ final} - RT_{initial} / RT_{medial/final}$). The relative effect of phoneme position on response time was significantly less for the ASD group (M = -0.08, SD = 0.54) compared to the control group (M = 0.22, SD = 0.25; z = 1.94, p = .03, one-tailed). The difference between groups of high IQ approached significance (ASD M = -0.10, SD = 0.57; control M = 0.20, SD = 0.26; z = 1.58, p = .06). No differences were found between groups of low IQ (ASD M = 0.04, SD = 0.41; control M = 0.33, SD = 0.18; z = 0.94, p = .18), or groups excluding individuals who had experienced delay in language development (ASD M = -0.05, SD = 0.56; control M = 0.21, SD = 0.26; z = 1.26, p = .11), although the small group sizes likely accounted for the lack of significant differences.

Correlations

There was no significant correlation between A' and age in the ASD group (r_s = -.27, p = .15) or the control group (r_s = -.10, p = .59). In contrast, a positive correlation was found between A' and MA, which was significant in the control group (r_s = .41, p = .02), and approached significance in the ASD group (r_s = .32, p = .08). Accordingly, IQ was strongly associated to A' in the control group (FIQ r_s = .48; FIQ-BD r_s = .52, PIQ r_s = .44, VIQ r_s = .40, all p < .03) and in the ASD group (FIQ r_s = .42, FIQ-BD r_s = .54, PIQ r_s = .43, all p < .03; with the exception of VIQ r_s = .30, p = .10).

The relative effect of phoneme position on response time showed no reliable correlation with age in the ASD (r_s = -.13, p = .49) or control groups (r_s = -.14, p = .47). This index was also not related to MA in either group (ASD r_s = .03, p = .87, control r_s = -.20, p = .28; no difference between coefficients, z_{r1-r2} = 0.86, p = .39). Furthermore, no association between the relative effect of phoneme position and IQ was found in either group (ASD r_s = .16 to .27, all p > .15; control r_s = -.10 to .27, all p > .16).

6.2.5 Summary of Phoneme Segmentation Task

Despite the prediction that the majority of TD participants were expected to have developed general phonological skills, discrimination accuracy on the Phoneme Segmentation Task improved with age, particularly in female participants. Furthermore, the ability to discriminate between target absent and target present trials showed a positive association with IQ. Younger participants were slower to respond than older participants, but the relative effect of embedding the phoneme (i.e., longer times to determine the presence of the phoneme when placed as the medial/final sound, versus the initial sound) was equivalent across all age groups.

Contrary to predictions of weak central coherence, the ability to detect the presence of a target phoneme was significantly poorer in the individuals with ASD (regardless of language history) relative to their matched control group. Ability to discriminate between target present and target absent trials did not relate to age in the ASD and control groups, although it was strongly associated with IQ and MA.

In line with predictions from weak central coherence theory, however, the ASD group showed reduced effects of the embedding context. While control and TD participants took longer to judge the presence of a phoneme when embedded as the medial or final sound, relative to the initial sound, no difference in response time between the two positions was observed in the ASD group. Since timing was from the end of each word, actual time from the target phoneme is underestimated, especially for initial positions, hence differences in response time for initial versus medial/final is underestimated for all groups.

Two candidate variables were taken from the Phoneme Segmentation Task and used in later within- and between-domain analyses: the non-parametric measure of discrimination accuracy (A') and the relative effect of phoneme position on response time. As weak central coherence predicted superior ability to disembed the target phoneme, good performance on trials when the phoneme was placed in the medial or final position (12/15 or more correct) was used to classify individuals as showing weak coherence.

6.3 Chord Segmentation Task

The hierarchical structure of tonal music is well documented in the study of perception and cognition of music. Tillman and Bigand (2004) define the Western tonal musical system as a "hierarchical grammar based on a restricted set of twelve *tones* ... These tones appear together in groups of three tones that define *chords* and create a second-order level of the musical system" (p. 212). Although certain chords are used together to form keys and cohesive harmonies, musical chords themselves have strong gestalt properties. These groups of notes are typically perceived as single objects, and it can be difficult to determine whether a particular tone has been heard in a chord. An analogy can be made in the visual domain with the EFT, which requires the ability to locate a simple figure embedded within a more complex gestalt form. A bias towards featural processing may result in an enhanced ability to separate and identify a single note from its context (Heaton, 2003). The ability to segment a musical chord was therefore selected as a marker of weak coherence in the auditory-perceptual domain.

Chord segmentation in ASD

Heaton, Pring and Hermelin (1999, Experiment 3) included a measure of chord segmentation in their case study of Dominic, a musically untrained boy with autism who demonstrated exceptional musical perception. The task involved listening to chords that were each followed by individual notes. A judgment was required as to whether or not the individual note was part of the chord that preceded it. Dominic showed superior performance on this task and correctly categorised 94% of the 48 target notes. Within a group of 13 children with autism (7 to 15 years), four children showed excellent chord segmentation skills, categorising at least 80% of the notes correctly. However the remaining children with autism performed at a similar level to age- and intelligence-matched control groups who categorised correctly, on average, 59% of the target notes. Heaton et al. (1999) surmised that Dominic's superior ability was due to a tendency to focus on the local rather than global aspects of musical stimuli. The study did not provide evidence however, for an enhanced ability to detect the individual elements of chords amongst the majority of children with autism.

Heaton (2003) investigated segmentation ability in the music domain in a subsequent study of children with autism and age- and intelligence-matched controls (7 to 15 years). After being familiarised with four target notes, the children with autism were superior to control children in identifying which of the four notes was absent when only three were played together as a chord. Without the prior familiarisation of target notes, the disembedding ability of the two groups was equivalent. This was measured by the Chord Segmentation Task described in Heaton et al. (1999) where participants determined whether a single note was part of a previously heard chord. Heaton (2003) concluded that superior pitch memory was able to facilitate or bias the children with autism towards a segmentation strategy. When the target notes were not familiar, the children with autism were shown to succumb to the holistic properties of musical chords.

Chord segmentation in typical development

Age-related improvements have been demonstrated on tasks measuring musical ability (e.g., Gromko & Poorman, 1998; Nelson & Barresi, 1989; Rowntree, 1973; S. Taylor, 1973). Standard measures of musical ability often include a measure of chord segmentation. For example, in the *Chords* subtest from the Bentley Test (Bentley, 1985) participants hear a group of notes played simultaneously and are required to state the number of notes they hear, with the actual number ranging from two to four. The *Chord Analysis* subtest from Wing's Standardised Tests of Musical Intelligence (H. Wing, 1968) is a similar detection task, but the number of notes presented in a single chord ranges from one to six. Both tasks have been found to relate to general intelligence in children (Lynn & Gault, 1986; Lynn, Wilson, & Gault, 1989). In a group of 10-year-olds (N = 217), Lynn et al. found the Chords subtest had a factor loading of 0.45 on Spearman's *g*. Lynn and Gault tested 49 boys and 44 girls (9 to 11 years) on the Chord Analysis subtest, and found this measure correlated significantly with the Raven's Standard Progressive Matrices (r = .27,

$p < .05$) as well as a measure of verbal reasoning ability. No gender differences in music aptitude were detected in this study.

Although gender differences in electrical brain responses have been detected while processing chord sequences in young children (Koelsch, Maess, Grossmann, & Friederici, 2003), to the author's knowledge, no investigation of single chord segmentation skill comparing males and females has been reported in the literature.

Predictions

Although superior disembedding of musical chords was not found to be pervasive in children with autism (Heaton, 2003; Heaton et al., 1999), this ability may still define a subgroup of individuals with weak central coherence. The Chord Segmentation Task developed by Heaton et al. (1999) was therefore incorporated in the present study to assess individual differences in local processing of musical stimuli in TD and in ASD. Participants were asked to state whether a subsequently played note had been one of the component notes in a chord.

It was predicted that weak coherence would be shown by an enhanced ability to mentally segment musical chords in order to identify individual components. Strong coherence, in contrast, would be demonstrated by a difficulty to overcome the gestalt properties of a chord in order to detect the presence of a note. As the literature has provided mixed results in the local-global processing of music stimuli in ASD, a conservative approach was taken and two-tailed tests were used for all musical processing tasks reported in this chapter.

According to the literature on typical development, music aptitude is reported to increase with maturity and intelligence as well as degree of musical training and experience (Bever & Chiarello, 1974). The ability to segment a chord into its component parts is therefore predicted to improve in accordance with these variables. Individual differences in the ability to identify a single note from a chord, over and above individual differences in development and music training, are also predicted in line with central coherence theory.

6.3.2 Method

Materials

Sixteen musical chords (8 major and 8 minor), randomised across keys, were used. The chords were tonic triads comprising the first, third and fifth degrees of the scales of the given keys. Comparison notes were the first or fifth degrees of the scales (8 stimuli correctly judged as yes) or the second or fourth degrees of the scales (8 stimuli correctly judged as no). In contrast to Heaton (2003), comparison notes that were the third (correct) and sixth (incorrect) degrees of the scale were not included. These notes were previously reported to be the easiest to judge correctly (Heaton, personal communication) and were removed to shorten the duration of the task.

Chords and single notes were recorded from a Casiotone 202 electronic keyboard (acoustic piano setting) by a Sony Minidisk recorder. The stimuli were played to participants through the Minidisk player using high quality speakers. The set of chords, listed in order of presentation, is shown in Appendix I.

Procedure

The training phase was adapted from Heaton (2003) to be more age-appropriate for the present age range. Participants were told to listen to three sequentially played notes (C, E, G) and then to listen to how they sounded when played simultaneously (C major chord). The experimenter gave the following instructions: "*On this task you will hear three notes together, and then one on its own. I want you to tell me whether the single note was part of the group you just heard.*"

One practice item was presented (C major chord, comparison note C) and corrective feedback was provided. The experimental phase consisted of 16 trials, with feedback given on the first two trials only. No time limit was imposed on participants to provide a response. If requested, the participant was able to hear the chord and comparison note again. One repetition of each trial was permissible, but this was limited to a maximum of three trials during the task. The experimenter recorded all responses on a score sheet.

6.3.3 Typical Development Results

All TD participants were administered the Chord Segmentation Task. Signal detection theory was used to measure participants' sensitivity to the presence and absence of target notes embedded in the chords (A', Grier, 1971) and response bias (B", Donaldson, 1992). Table 6.5 presents mean accuracy scores and A' values for the four age groups. As data were significantly negatively skewed (z-scores ranged from -3.6 to -1.4), and tests of normality were failed (Kolmogorov-Smirnov, all $p < .02$), non-parametric statistical tests were used.

Age groups were equivalent on the bias index, a measure of an individuals' tendency to favour a positive over a negative response ($\chi^2 = 6.84$, $p = .08$). Mean bias scores were negative for all groups revealing a tendency for participants to say the target note was present in the chord (range -0.25 to -0.02). Male and female participants were also equivalent on the bias index when comparing between and within age groups (all $p > .13$).

As shown in Table 6.5, mean values of A' were low across age groups (all $\leq .70$) indicating that participants had difficulty distinguishing whether a note was present or absent. One-sample t-tests were performed to verify if mean performance differed from chance (4 out of 8 correct). All age groups performed above chance when identifying notes that were present in the chords (all $t > 3.90$, $p < .0005$). The 8-10 year group did not show mean performance that differed from chance when identifying the absence of a note ($t_{(53)} = 1.15$, $p = .26$). All older groups performed above random performance levels (all $t > 2.13$, $p < .04$).

Chance levels of performance were also determined at an individual level by conducting separate binomial tests. Performance was above chance for 15% of 8-10 year group, 30% of 11-13 year group, 43% of 14-16 year group and 32% of 17-25 year group. The distribution of these frequencies was significantly different across groups ($\chi^2 = 10.26$, $p = .02$). Pairwise comparisons revealed that the youngest age group was significantly different from the two oldest age groups (14-16 years $\chi^2 = 10.30$, $p = .001$; 17-25 $\chi^2 = 4.57$, $p = .03$).

Table 6.5 Number of correct responses on target present and target absent trials and the corresponding measure of sensitivity (A') for each age group on the Chord Segmentation Task: Mean (*SD*)

Age group (years)	N	Correct detections (max = 8)	Correct rejections (max = 8)[a]	A'[b]
8-10	54	4.81 (1.53)	3.74 (1.66)	0.50 (0.28)
11-13	43	5.19 (1.55)	4.56 (1.71)	0.61 (0.28)
14-16	51	5.39 (1.59)	5.35 (1.62)	0.70 (0.24)
17-25	56	5.46 (1.58)	4.80 (1.92)	0.66 (0.28)

[a] 8-10 < 11-13, 14-16, 17-25, all $p < .05$; 11-13 < 14-16, $p = .02$
[b] 8-10 < 11-13, 14-16, 17-25, all $p < .04$

A significant effect of age group was found on A' ($\chi^2 = 17.0$, $p = .001$), with the youngest age group performing significantly lower than all older age groups (all $z > 2.00$, $p < .04$). The age-related improvement on the Chord Segmentation Task was more apparent in females than males (see Figure 6.4). Significant effects of age group were found on A' for female participants ($\chi^2 = 15.24$, $p = .002$), but not males ($\chi^2 = 6.31$, $p = .10$). Females in the youngest age group had significantly lower A' values than all older age groups (all $z > 2.11$, $p < .03$).

Significant effects of age group were found on judging the absence of a note ($\chi^2 = 20.9$, $p < .0005$) but not for identifying the presence of a note ($\chi^2 = 5.0$, $p = .17$; see Table 6.5). When analysing by gender, significant age group effects correctly judging the absence of a target note were found for females ($\chi^2 = 16.17$, $p = .001$), but not males

($\chi^2 = 7.12, p = .07$). Females in the youngest group made fewer correct rejections than all older age groups (all $z > 2.44, p < .02$).

No main effects of gender were found for number of correct identifications ($z = 1.03$, $p = .30$), correct rejections ($z = 0.17, p = .87$), nor A' ($z = 0.36, p = .72$). Male and female participants also performed at an equivalent level within each age group (all $z < 1.67$, $p > .09$).

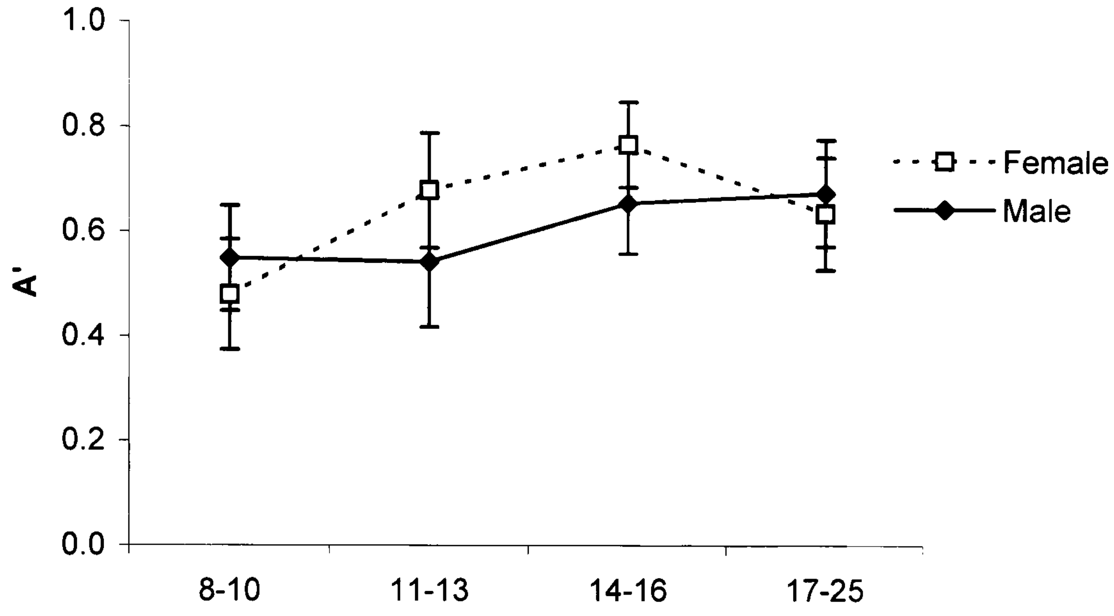


Figure 6.4 Sensitivity index (A') as a function of age group for male and female participants on the Chord Segmentation Task. Error bars show 95% confidence intervals.

Correctly identifying the presence of a target note was relatively easier than correctly identifying the absence of a target note over the whole TD sample (Wilcoxon's Signed Ranks Test, $z = 4.69, p < .0005$). This was pattern was found in all age groups (all $z > 2.30$, $p < .03$), except for the 14-16 year group who performed equivalently in the two conditions ($z = 0.34, p = .73$).

Correlations

The ability to judge the presence and absence of target notes (A') significantly increased with age in the TD sample ($r_s = .26$, $p < .0005$; males $r_s = .26$, $p = .01$; females $r_s = .29$, $p = .003$). Task performance was also related to overall ability, with significant correlations found between A' and all IQ measures (FIQ $r_s = .23$, $p = .001$, FIQ-BD $r_s = .19$, $p = .006$, VIQ $r_s = .18$, $p = .01$, PIQ $r = .22$, $p = .002$). Correlations between A' and IQ were generally stronger in females (FIQ $r_s = .25, p = .01$, FIQ-BD $r_s = .22, p = .02$, VIQ $r_s = .20$, $p = .04$, PIQ $r = .25$, $p = .01$) than that in males (FIQ $r_s = .21$, $p = .04$, FIQ-BD $r_s = .17, p = .10$, VIQ $r_s = .17, p = .10$, PIQ $r = .19, p = .06$), but did not reliably differ in magnitude (all $z_{r1-r2} < 0.46$, $p > .64$). Music experience was expected to relate to Chord Segmentation performance and this was confirmed by significant positive correlations between A' and the mean number of years of music training (r_s = .40; males r_s = .45; females r_s = .36, all p < .0005).

6.3.4 *ASD and Control Results*

The Chord Segmentation Task was not administered to one 14-year-old boy from the ASD group who had refused to attend the final session. The groups remained well matched on age and IQ after the exclusion of this participant from analyses.

Mean accuracy scores and measure of sensitivity (A') on the Chord Segmentation Task for the ASD and control groups are presented in Table 6.6. As the data from each group failed tests of normality (p < .02), non-parametric statistical tests were used.

The index of task performance was low for all groups (mean A' < .65), indicating an overall difficulty in determining the presence versus absence of a note. Mean accuracy scores for both groups were, however, above chance levels of responding. One-sample t-tests verified that overall performance was significantly different from chance (ASD $t_{(29)}$ = 3.19, p = .003; control $t_{(30)}$ = 3.46, p = .002). This was apparent for both groups when correctly identifying the presence of target notes (all t > 3.6, p < .001). Mean performance for correctly detecting the absence of notes was above chance performance for the control group ($t_{(30)}$ = 2.36, p = .03) but not for the ASD group ($t_{(29)}$ = 1.30, p = .20). It was investigated whether this finding had arisen from a bias towards reporting the presence of the note in the ASD group. The mean bias index (B'') for the ASD group was -0.07 (SD = 0.46), indicating a neutral response pattern, which did not differ significantly from the control group (M = -0.14, SD = 0.48; z = 0.59, p = .56). It could not be concluded therefore that a bias towards affirmative responses resulted in below chance performance on target absent trials in the ASD group.

Separate binomial tests were performed to determine chance levels of performance for individuals over the whole task. Groups were equivalent in the number of participants performing above chance: 33% of the ASD participants (10 from 30) and 32% of the control participants (10 from 31). Participants who performed above and below chance on the Chord Segmentation Task were equivalent in age and IQ within each group (all z < 0.86, p > .38).

No group differences were detected on A' (z = 0.04, p = .97, two-tailed). Furthermore, groups did not differ on correctly judging the presence of target notes (z = 0.13, p = .89) or their absence (z = 0.26, p = .79). The same pattern of results was found when group comparisons included participants with high IQ (all z < 0.52, p > .60) or low IQ only (all z < 0.32, p > .30).

Table 6.6 Number of correct responses on target present and target absent trials and the corresponding measure of sensitivity (A') for each group on the Chord Segmentation Task: Mean (*SD*)

	Group	N	Correct detections (max = 8)	Correct rejections (max = 8)	A'
All	ASD	30	5.23 (1.55)	4.53 (2.24)	0.62 (0.28)
	Control	31	5.13 (1.75)	4.71 (1.68)	0.61 (0.30)
High IQ	ASD	24	5.21 (1.44)	4.63 (2.10)	0.63 (0.26)
	Control	25	5.36 (1.55)	4.68 (1.77)	0.64 (0.29)
Low IQ	ASD	6	5.33 (2.07)	4.17 (2.93)	0.60 (0.36)
	Control	6	4.17 (2.32)	4.83 (1.33)	0.52 (0.34)

Although the number of correct detections of target notes was higher than the number of correct rejections, the difference between these two conditions did not differ in the ASD group (Wilcoxon's Signed Ranks Test, $z = 1.29$, $p = .20$, two-tailed), or in the control group ($z = 1.26$, $p = .21$). This was in contrast to the TD sample, where correct identifications of the presence of a note were easier than identifying its absence.

In order to determine whether subgroups of participants with exceptional performance on the Chord Segmentation Task could be found in the ASD and control groups (as in Heaton et al., 1999), the frequency distribution for the total number of correct judgments was examined (see Figure 6.5). Two ASD participants produced perfect scores, which did not occur in the control group (although one control participant made only one error over the 16 trials). Overall, however, no obvious difference in the distribution of scores between groups was evident.

Eight participants, four each from the ASD and control groups, scored high overall (at least 14 correct from 16) and their group characteristics were explored further. The age range of the four ASD participants (13.8 to 16.6 years) was lower than the four control participants (16.0 to 18.4 years), but not at a statistical significant level ($z = 1.44$, $p = .15$). However, the ASD participants had significantly lower MA ($M = 11.1$ years, $SD = 3.4$) than the control participants ($M = 16.0$ years, $SD = 2.7$; $z = 2.02$, $p = .04$). Three ASD participants were of low general ability (FIQ ranged from 52 to 78), and one had normal IQ (FIQ of 98). In the control group, one low-functioning individual was identified (FIQ of 68), while the remainder were of high ability (FIQ ranged from 103 to 106). A trend was apparent for the ASD participants to have lower VIQ than the control participants ($z = 1.89$, $p = .06$), but did not vary significantly on all other IQ measures (all $z < 1.74$,

$p > .07$). Furthermore, all control participants had received formal music training, of at least one year (maximum 10 years). In contrast, only two of the ASD participants had received any music training (one to two years). The difference in years of music training did not reach statistical significance ($z = 1.76, p = .08$, two-tailed).

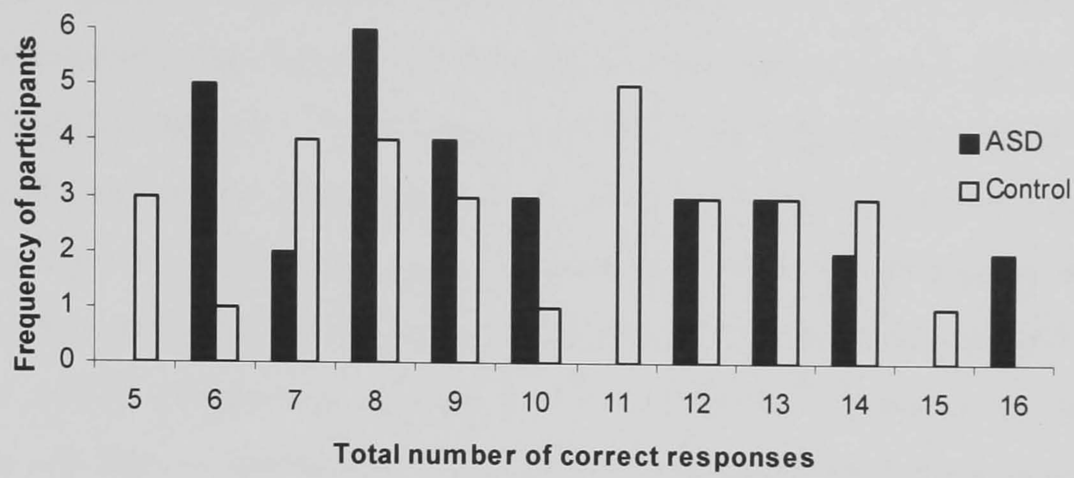


Figure 6.5 Frequency distribution of the total number of correct responses (max = 16) on the Chord Segmentation Task for ASD and control participants.

Correlations

The ability to determine the presence and absence of target notes within chords did not correlate significantly with age in the ASD group ($r_s = -.01, p = .98$) or the control group ($r_s = .16, p = .39$). Taking developmental differences into account did not alter this relationship, with no reliable correlation shown between MA and A' (ASD $r_s = -.11$, $p = .57$; control $r_s = .19, p = .32$). Chord segmentation skill did not correlate with any IQ measure in the ASD group ($r_s = -.007$ to $-.11$, all $p > .58$). The correlation coefficients between the IQ indices and A' in the control group were of similar magnitude to the TD sample, but none reached statistical significance ($r_s = .16$ to $.25$, all $p > .17$). Performance on the Chord Segmentation Task did not correlate significantly with the number of years of music training in the ASD group ($r_s = .19, p = .31$) or the control group ($r_s = .29, p = .11$).

6.3.5 Summary of Chord Segmentation Task

It can be concluded from the Chord Segmentation Task that identifying individual elements in musical configurations such as chords is very difficult for typical populations as well as for individuals with ASD. This ability showed significant improvements with age in the TD sample, most notably after the age of 8-10 years and in female participants. Furthermore, IQ and formal music training had significant influences on task performance in the TD sample.

Participants in the ASD group did not show superior disembedding of music chord stimuli as predicted by weak central coherence theory. Presenting the candidate note *before* the chord might be more analogous to the EFT, however, and might have resulted in better performance in the ASD group. The characteristics of a small subgroup of participants who did perform well on the Chord Segmentation Task were contrasted between groups. Of interest was the observation that the ASD participants were of significantly younger MA and had generally fewer years of musical training than the control participants.

TD participants were generally better at determining the presence rather than absence of notes embedded within chords. The converse was found by Heaton (2003), where both ASD and control children (7 to 14 years) showed better identification of the absence of notes, than their presence. Heaton proposed that determining the absence of a note might reflect an ability to identify notes that sound discordant to the previously heard chord, rather than disembedding skill per se. This suggests that the correct identification of notes embedded within chords may provide a more adequate assessment of weak central coherence. In light of this finding, two candidate variables were taken from the Chord Segmentation Task as continuous measures of performance: sensitivity of discrimination overall (A') and the number of correct detections of the embedded note. Participants were categorised as showing weak central coherence if they showed good performance on target present trials (at least 6 from 8 correct).

6.4 Pitch Identification Task

Absolute pitch refers to the ability to produce or identify the pitch of a sound without any external reference point (Takeuchi & Hulse, 1993). It is a rare phenomenon; most individuals typically perceive the "melodic and harmonic relations among pitches instead of the absolute pitches themselves" (Takeuchi & Hulse, 1993, p. 345). Absolute pitch has an estimated incidence of 1 in 10,000 in the general population (Profita, Bidder, Optiz, & Reynolds, 1988), although the prevalence may depend on how absolute pitch is conceptualised. Levitin (1994), for example, considers that absolute pitch should be viewed as two distinct components: *pitch memory*, "the ability to maintain stable, long-term representations of specific pitches in memory, and to access them when required" and *pitch labelling*, "the ability to attach meaningful labels to these pitches" (p. 415). Implicit memory for pitch has been found to occur more frequently than previously thought (Levitin, 1994; Schellenberg & Trehub, 2003), whereas the explicit ability to associate a label to a specific pitch is rare.

As not all proficient musicians possess absolute pitch, it cannot be used as a reliable marker for outstanding musical talent in typical populations. Absolute pitch is however, prevalent in all musical savants that have been documented in the literature (Heaton & Wallace, 2004). A relatively high prevalence of absolute pitch has also been found in ASD (Heaton, Hermelin, & Pring, 1998). It was proposed that a cognitive style that biases processing towards local features might predispose individuals with autism towards absolute pitch ability. An enhanced memory for pitch may therefore be a demonstration of weak coherence when processing musical stimuli. The ability to identify this *absolute* property of single notes was considered a plausible indicator of local processing bias in the present study.

Absolute pitch in typical development

Much research has been concerned with pitch processing and pitch memory in typically developing populations (for reviews see Takeuchi & Hulse, 1993; Trehub, 2003). It is known that infants' and children's performance in detecting and discriminating between sounds is often poorer than that of adults (Maxon & Hochberg, 1982). This is said to reflect the maturation of the primary auditory pathway and specific central processes which continue into childhood and influence hearing (L. A. Werner, 1996). Developmental effects have also been reported in the duration of memory for tone pitch. Keller and Cowan (1994) compared children aged 6 to 7 years, 10 to 12 years, and adults, and found the retention of sensory pitch information significantly increased with age. The authors concluded that this developmental effect was not due to differences in memory retention strategies or ability to discriminate between tones.

It is a puzzling question as to why absolute pitch develops in some individuals and not in others. Zatorre (2003) provides a review of studies looking at the role of genetic factors in absolute pitch, as well as neurological differences that have been investigated through functional and structural neuroimaging studies. There is much evidence that an environmental influence, in the form of very early musical training, contributes to the development of absolute pitch. All infants are said to possess absolute pitch but a shift to relative pitch processing typically occurs in early childhood (Saffran & Griepentrog, 2001; Takeuchi & Hulse, 1993). Absolute pitch can be retained however, if the child is exposed to musical training during a critical period of development, approximately between the ages of five and seven (Saffran & Griepentrog, 2001). Exposure to early music training does not necessarily result in absolute pitch ability, however. Baharloo, Service, Risch, Gitschier and Freimer (2000), for example, found that individuals with absolute pitch (N = 4) in their sample of over 600 music students all began their music training before 6 years of age, but the overwhelming majority of people exposed to early music lessons did not have absolute pitch (N = 135 from 139). Takeuchi and Hulse (1993) speculate that the nature of early music training is critical and absolute pitch may only develop if the training includes the association of pitch names to isolated pitches. It is also likely that individuals differ in their potential to develop absolute pitch.

A study by Costa-Giomi, Gilmour, Siddell and Lefebvre (2001) lends support to the possibility that a difference in cognitive style also contributes to the development, or is at least an associated factor, of absolute pitch. They found that musicians who possessed absolute pitch scored significantly higher than musicians without absolute pitch on the Hidden Figures Test (akin to the EFT). As a local processing style is beneficial to performance on the Hidden Figures Test, this finding provides some evidence for weak coherence operating across different processing domains. Although musicians with and without absolute pitch were not equated for overall intelligence, the two groups were equivalent on two further measures of visual perception requiring spatial visualisation (identifying objects in different positions) and global coherence (identifying objects from a sequence of disarranged segments).

Absolute pitch in ASD

Superior detail processing has been observed in autism within the musical domain. Applebaum, Egel, Koegel and Imhoff (1979) for example, compared three musically naïve children with autism and three musically-experienced TD children on their ability to sing back notes and musical pieces. They found that the children with autism performed as well or better than the more intelligent and musically experienced controls. Heaton et al. (1998) commented that this demonstration of 'musical echolalia' in the children with autism should be distinguished from absolute pitch ability, where stable long-term representations of individual pitches are made.

More recent experimental studies have confirmed that pitch memory and discrimination is superior in autism (Bonnel et al., 2003; Heaton, 2003; Heaton et al., 1998, 1999; Mottron et al., 2000). In one study Heaton et al. (1998) paired single notes with animal pictures as a simplified way to test for stable note-name mappings, necessary for absolute pitch. Four pitch-animal pairings were presented 12 times in random order. In comparison to MA-matched control participants, musically naïve children with autism (age range 7 to 13 years, IQ range 55 to 127) were superior in identifying the animal associated with each single note. The children with autism were still able to remember the associations one week after being exposed to the pairings.

The Pitch Identification Task from Heaton et al. (1998) was selected for the present study with two modifications made to the original. First, the training phase was shortened so that participants were exposed to each pitch-animal pair six times. Secondly, participants were tested on their learned associations on one occasion, one minute after the training phase. This delay duration was selected as studies have shown individuals who possess absolute pitch have little memory decay of pitches, but in the absence of absolute pitch recall deteriorates to chance levels of performance following a retention interval of one minute or more (Krumhansl, 2000; Takeuchi & Hulse, 1993). Both modifications were employed to shorten the task overall and maintain motivation levels.

6.4.2 Line Orientation Task

To compare memory for exact properties of a stimulus, and ascertain whether music is a special domain, an analogous task to Pitch Identification was made for the visual modality. As mentioned above, absolute pitch is found in all musical savants reported in the literature and it may be an important mediating or associated factor for musical talent. Another domain where savant skills are often observed is that of perspective drawing (Hermelin & O'Connor, 1990; Mottron & Belleville, 1993; O'Connor & Hermelin, 1987; Selfe, 1977). Mottron and Belleville (1993), for example, in their case study of a 34-year-old male with Asperger syndrome and exceptional graphic abilities, describe the "extreme precision he shows for graphic details such as angles, curves and lines" (p. 303). This led to the hypothesis that an enhanced ability to discriminate line angle may be an associated feature of exceptional artistic talent. In a similar vein to pitch processing, making relative judgments of the orientation of a line is much easier than producing or identifying the absolute orientation without any external reference point. As a visual analogy to the Pitch Identification Task, the Line Orientation Task was developed to assess the ability to make a stable or absolute representation of line angle in the absence of an external reference. It was postulated that this ability might result from a perceptual bias towards featural processing or weak central coherence.

Task Development

In the Line Orientation Task participants were required to learn an association between animal pictures and lines of particular orientation. Two conditions of the task were initially piloted where the four target line angles differed either by 4° (26°, 30°, 34° and 38°) or 5° (20°, 25°, 30° and 35°). These divisions were selected on the basis of discrimination thresholds for line orientation obtained by Chen and Levi (1996). The authors determined that the 'just noticeable difference' in line orientation for oblique lines (i.e., not at vertical or horizontal orientations) varied between 0.5° and 4.0° for young adults.

Two design features were made to prevent the participant using external visual cues to determine each angle. First, all vertical and horizontal lines near the line stimulus were removed by covering the touch screen with a piece of felt. The participant viewed the screen through an oval hole cut into the material. Secondly, each line angle appeared to constantly change in length (i.e., starting from 6cm, expanding to 9cm, decreasing to 3cm, returning to 6cm). This design prevented participants from using the vertical separation between the two line endpoints as a cue to "measure" each angle. The ability to form a stable absolute representation of each angle could therefore be assessed for each participant.

A total of 28 children, ranging in age from 8 to 15 years, were administered either the 4° or 5° pilot of the Line Orientation Task. Accuracy rates were low for both conditions with the mean number of correctly identified line-animal pairs not differing from chance. As a consequence, a third condition was constructed where the lines differed in angle by 6° (26°, 32°, 38° and 44°). This condition was piloted on 25 children ranging in age from 7 to 15 years. The 6° of line separation produced a greater range of accuracy (6% to 81% of items correct) with mean performance above chance levels (M = 36%, SD = 20%). The four lines separated by 6° (see Figure 6.6) were therefore selected as the test stimuli for the Line Orientation Task in order to detect individual differences in performance, while avoiding floor and ceiling effects.

6.4.3 Symbol Task

The Line Orientation and Pitch Identification Tasks can essentially be viewed as paired associate learning tasks using test stimuli that vary along a single dimension (pitch tone and line angle). In order to conclude that poor performance on the experimental tasks was due to a specific deficit in forming an absolute representation of pitch and/or line angle, and not a difficulty in associative learning per se, a control task was designed that employed discrete test stimuli that did not differ along a continuum nor benefit from a local processing bias.

The *Symbol Task* was developed as a control measure, in which four abstract designs were presented in conjunction with four different animal pictures (see Figure 6.6). In line with the administration of the Pitch Identification and Line Orientation Tasks, participants were tested on their ability to learn the association between a design and an animal picture. Although developmental differences in strategic memory processes may occur, the abstract designs were constructed to minimise the possibility of participants' forming semantic associations between the design and animal. An attempt was therefore made to obtain a 'pure' measure of associative learning for each participant.

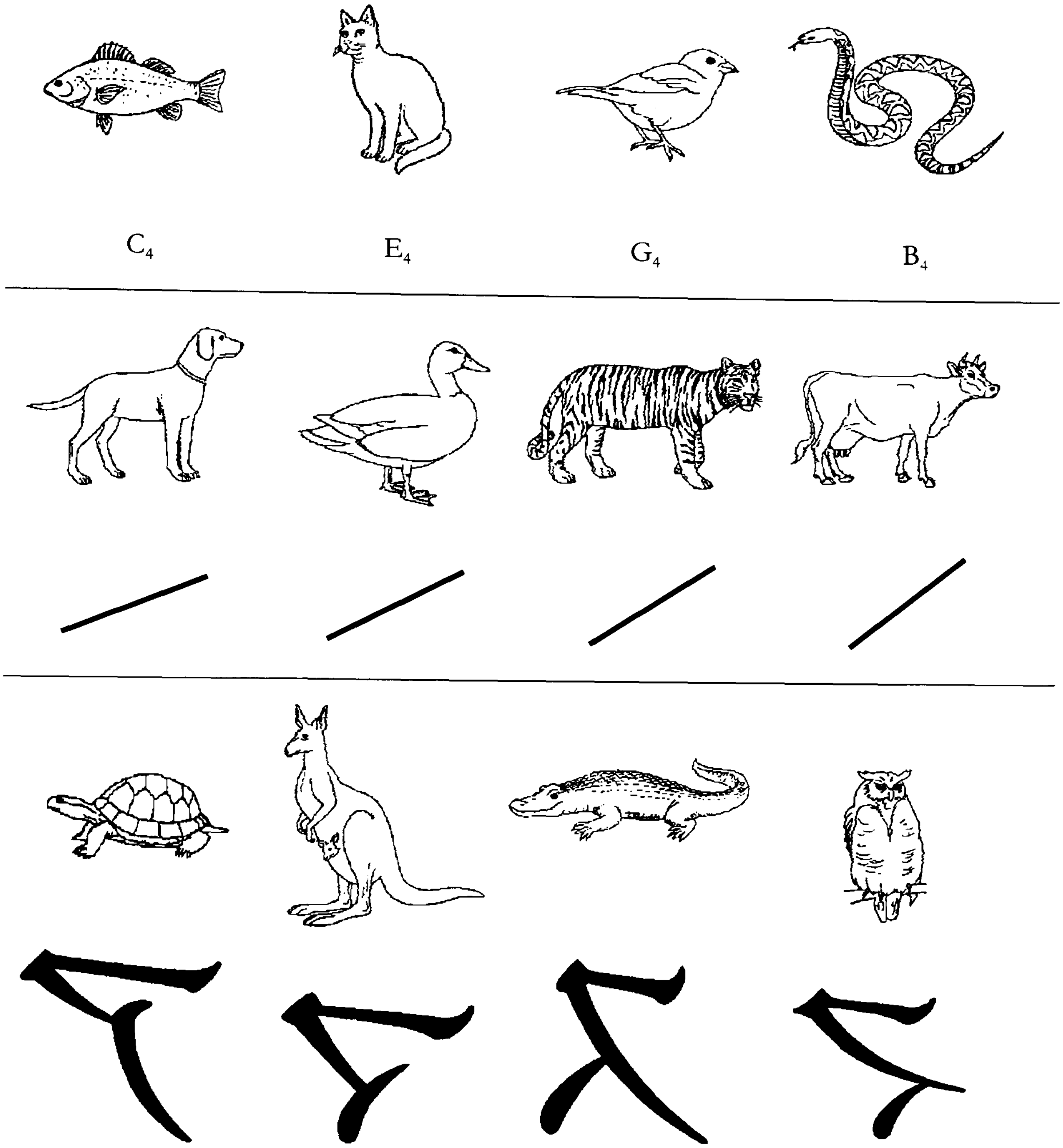


Figure 6.6 Pitch-animal pairs (top), line-animal pairs (middle) and symbol-animal pairs (bottom) used in the Pitch Identification, Line Orientation and Symbol Tasks respectively.

The Symbol Task was piloted on 21 children, ranging in age from 7 to 15 years. Accuracy rates were relatively high with the mean number of correctly identified symbol-animal pairs at 66% (*SD* = 28%). Performance was at, or close to, ceiling for 8 of the 21 participants (94% to 100%), while 5 participants performed below chance levels (19% to 38%). The pilot data confirmed that the Symbol Task could assess individual differences in paired associative learning and function as an appropriate control to the Line Orientation and Pitch Identification Tasks.

Predictions

Developmental differences in associative learning are often found in TD populations and have been attributed to corresponding age-related changes in strategy use (Kail, 1990; Kee, 1994). Improvement with age and ability are therefore expected on the three associative learning tasks. Individual differences in performance, over and above developmental effects, are predicted in line with central coherence theory for the Pitch Identification and Line Orientation Tasks. Individuals with a tendency towards weak central coherence are expected to focus on the individual pitch more than on the musical context in everyday experience of music and, as a consequence, demonstrate good retention and discrimination of individual tones on the Pitch Identification Task. Similarly, in the visual analogy to absolute pitch processing, accurate recall of absolute line orientation could be driven by a local processing style.

6.4.4 Method

Materials

Stimuli for the Pitch Identification Task consisted of four musical notes (C_4, E_4, G_4, and B_4) recorded from a Casiotone 202 electronic keyboard (acoustic piano setting). Each note was stored as an individual sound file. Lines drawn at four angles from the horizontal (26°, 32°, 38° and 44°) were used in the Line Orientation Task. For each angle a series of 21 bitmap files were created in which the length of the line varied from 3 cm to 9 cm. When played in succession, the line appeared to constantly change in length, while the angle remained static. Four abstract designs, constructed to be fairly meaningless, were used in the Symbol Task. The designs also shared some local features in order to make feature-animal mappings difficult. Four adult raters concurred that it was difficult to form semantic associations to each symbol.

Each of the notes, line angles and symbols were presented in conjunction with a picture of a different animal. All pictures were selected from the Snodgrass and Vanderwart (1980) Picture Norms Set. Each animal picture had a high percentage of name agreement (88% to 100%) and familiarity rating (mean scores 1.7 to 4.6 on a scale from 1 to 5). Furthermore, the mean name agreement and familiarity rating for the four animals did not differ significantly across the three tasks. Figure 6.6 presents each pairing of an animal with a note, line angle, or abstract symbol.

The presentation of all stimuli was controlled by SuperLab Pro software run by a laptop computer. Participants viewed the stimuli on a touch screen, which also served as the response input module.

Procedure

The Line Orientation Task was administered to all participants in the first test session, while the Pitch Identification Task took place in the second session and the Symbol Task in the final session. The method of administration of each task was comparable. Participants were first familiarised with the four animals that appeared together on the screen and were informed that each animal had a favourite sound/slope/symbol[4]. It was explained to the participant that they had to remember the sound/slope/symbol that went with each animal.

The participant was then exposed to each animal-stimulus pair in six blocks of trials, resulting in 24 training trials. The animal-stimulus pairs were presented in a fixed random order within each block. On each trial the animal picture appeared first (in the top left quadrant of the screen) and then the associated line angle or symbol (appearing in the bottom right quadrant) was exposed for five seconds. In the Pitch Identification Task each animal picture appeared centrally and remained on the screen for five seconds after the associated note was sounded. For each presentation the participant was told, "*this is the sound/slope/symbol that the [animal] likes best.*"

Each training block was separated by the words "*Let's see them again…*" appearing on the screen, while the last block was cued by "*And one last time….*" Participants were reminded to try their best to remember the sound/slope/symbol that went with each animal.

Following the training session the participant was engaged in conversation about an unrelated topic for one minute. The testing phase then began with all four animals appearing on the screen (in the top, bottom, left and right quadrants) and remaining in the same position throughout the testing phase. Each sound/slope/symbol was presented individually four times in a pseudo-random order, such that no two identical stimuli appeared consecutively. The participant was asked to touch the animal picture on the screen that they thought went with each sound/slope/symbol. Each line angle was presented in a static form, of four different lengths, for ease of administration. Participants had an unlimited time to respond. While the line and symbol remained on the screen until the participant responded, each note could only be played once on each trial. No feedback was given during the testing phase and participants were told that they could make a guess if they were not sure of the correct answer.

---

4 The wording was modified for younger participants: "symbol" was replaced by "picture" and "slope" was replaced by "favourite slope that they liked to slide down."

6.4.5 *Typical Development Results*

The means and standard deviations for the three associative learning tasks are shown in Table 6.7 for each age group in the TD sample. All participants were administered the tasks, however a data file from the Pitch Identification Task was lost for one 17-year-old female and her data were excluded on this task.

The mean number of correctly identified pitch-animal and symbol-animal pairs was negatively skewed in the three oldest age groups (z-scores -1.4 to -8.0) showing the tendency for scores to approach the maximum number correct. Conversely, the data were positively skewed in the youngest age group for identifying line-animal pairs (z-score = 2.1). As a result of the non-normality of the data, non-parametric statistical tests were used.

Table 6.7 Number of correctly identified items (maximum = 16)[†] on the three associative learning tasks for each age group: Mean (*SD*)

Age group (years)	N	Pitch[a]	Line[b]	Symbol[c]
8-10	54	8.43 (3.72)	4.89 (2.19)	10.02 (3.99)
11-13	43	10.91 (3.16)	6.42 (2.30)	12.86 (3.75)
14-16	51	11.94 (3.68)	7.59 (2.79)	13.61 (3.76)
17-25	56	11.05[*] (3.65)	8.32 (2.36)	14.61 (2.56)

[†]Chance performance = 4. [*]N = 55
[a]8-10 < 11-13, 14-16, 17-25, all $p < .001$
[b]8-10 < 11-13, 14-16, 17-25, all $p < .001$; 11-13 < 14-16, 17-25, all $p < .02$
[c]8-10 < 11-13, 14-16, 17-25, all $p < .0005$; 11-13 < 17-25, $p = .02$

Binominal tests were performed on the data to determine chance levels of performance within age groups ($p > .05$). The percentage of participants within each age group performing significantly above chance on each of the three associative learning tasks is presented in Figure 6.7. The distribution of participants with performance above chance was significantly different between age groups in each task (Line Orientation $\chi^2 = 44.16$, $p < .0005$; Pitch Identification $\chi^2 = 14.62$, $p = .002$; Symbol $\chi^2 = 9.47$, $p = .02$). Pairwise comparisons revealed that the number of participants in the youngest age group performing above chance on the Line Orientation Task and the Pitch Identification Task was significantly fewer than all older age groups (all $z > 2.50$, $p < .02$). The 11-13 year group also had significantly fewer participants performing above chance on the Line Orientation Task than the two oldest age groups (all $z > 2.37$, $p < .02$). On the Symbol

Task, only the youngest and oldest groups differed from each other in the relative proportion of participants performing above chance ($z = 2.91, p = .003$).

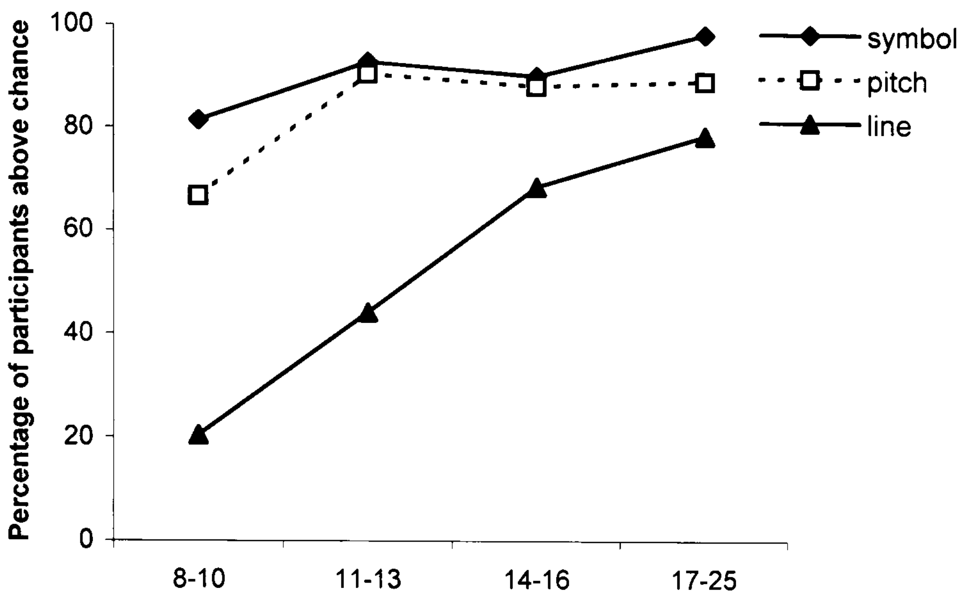


Figure 6.7 Percentage of participants in each age group who performed above chance in identifying stimuli-animal pairs.

Significant effects of age group were found in all three tasks: Pitch Identification, $\chi^2 = 26.3$, Line Orientation, $\chi^2 = 50.2$, and Symbol, $\chi^2 = 42.9$ (all $p < .0005$; see Table 6.7 and Figure 6.8). In all tasks, the number of stimulus-animal pairs correctly identified increased significantly after 8-10 years of age. Performance further improved after 11-13 years for identifying line-animal pairs, and symbol-animal pairs.

It was considered that participants might use relative processing to enable the correct identification of the two extreme stimuli on the Pitch Identification and Line Orientation Tasks (i.e., the highest/lowest sound, steepest/flattest slope), rather than the intended absolute processing of stimuli. When the two extreme stimuli were removed from analysis, significant effects of age group were found on both tasks ($\chi^2 > 21.5$, $p < .0005$) and confirmed that the same pattern of results were found for the two middle items as compared to all items. This suggested that the two tasks were tapping absolute processing of stimuli to some degree, and not purely relative processing.

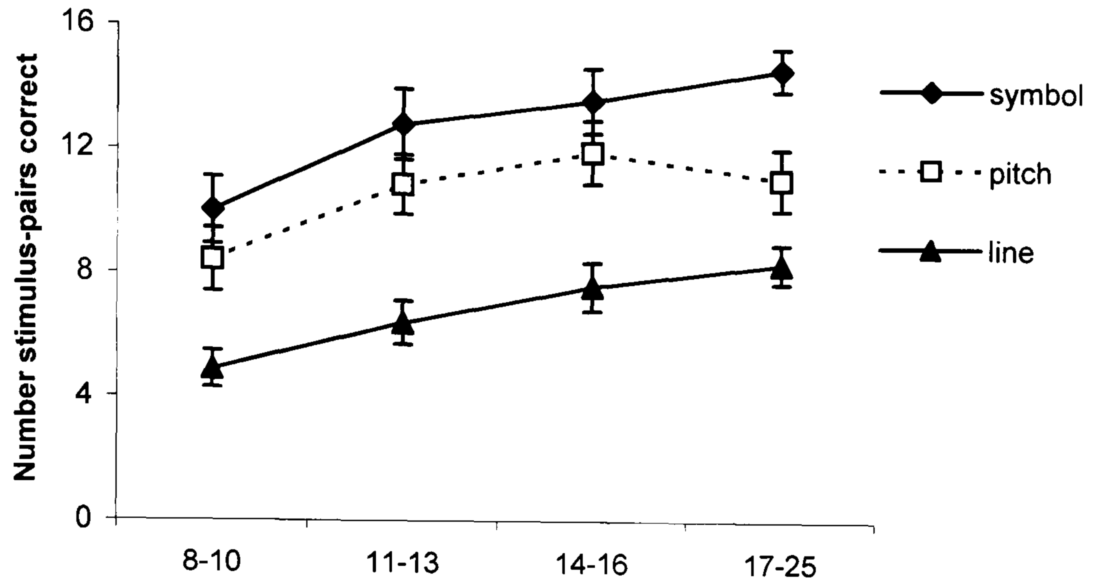


Figure 6.8 Mean number of correctly identified stimuli-animal pairings on the three associative learning tasks for each age group. Error bars show 95% confidence intervals.

Significant effects of age group were found for female participants on all three tasks (Line Orientation, $\chi^2 = 31.2$; Pitch Identification, $\chi^2 = 26.0$; Symbol, $\chi^2 = 39.9$, all $p < .0005$). For males, however, significant effects of age group were only found on the Line Orientation Task ($\chi^2 = 22.8$, $p < .0005$). Males in the youngest age group identified significantly fewer line-animal pairs than all older males (all $z > 1.96$, $p < .05$), while males in the oldest age group identified significantly more line-animal pairs than all younger males (all $z > 2.19$, $p < .03$).

The stronger effect of age group for females compared to males on the Pitch Identification and Symbol Tasks is depicted in Figure 6.9. An incongruous finding is the high performance of the 14-16 year females on the Pitch Identification Task, which was significantly higher than both the 17-25 year females ($z = 2.87$, $p = .004$) and the 11-13 year females ($z = 2.73$, $p = .006$). It was considered whether the influence of musical training on Pitch Identification could explain this finding. Females in the 14-16 year group had received more years of musical training on average than both the 11-13 and 17-25 year groups, although neither difference reached conventional levels of significance (see Appendix E). The analyses were repeated after removing participants who reported having more than one year of musical training. This excluded the majority of females in the 14-16 year age group (N = 16), but somewhat surprisingly, the remaining females (N = 5) still exceeded the oldest age group in Pitch Identification performance ($M = 13.40$, $SD = 0.89$, versus $M = 8.94$, $SD = 3.79$; $z = 2.54$, $p = .01$). This difference could not be explained by IQ differences between the two groups (all $z < 1.58$, $p > .11$).

Main effects of gender were found on the Symbol Task ($z = 1.97$, $p = .05$), with males surpassing females overall in number of symbol-animal pairs correctly identified. When analysing by age group, this effect was only significant in the youngest age group (z = 2.63, p = .009). No overall difference between male and female participants was found on the Line Orientation (z = 1.29, p = .19), or Pitch Identification Tasks (z = 1.32, p = .19).

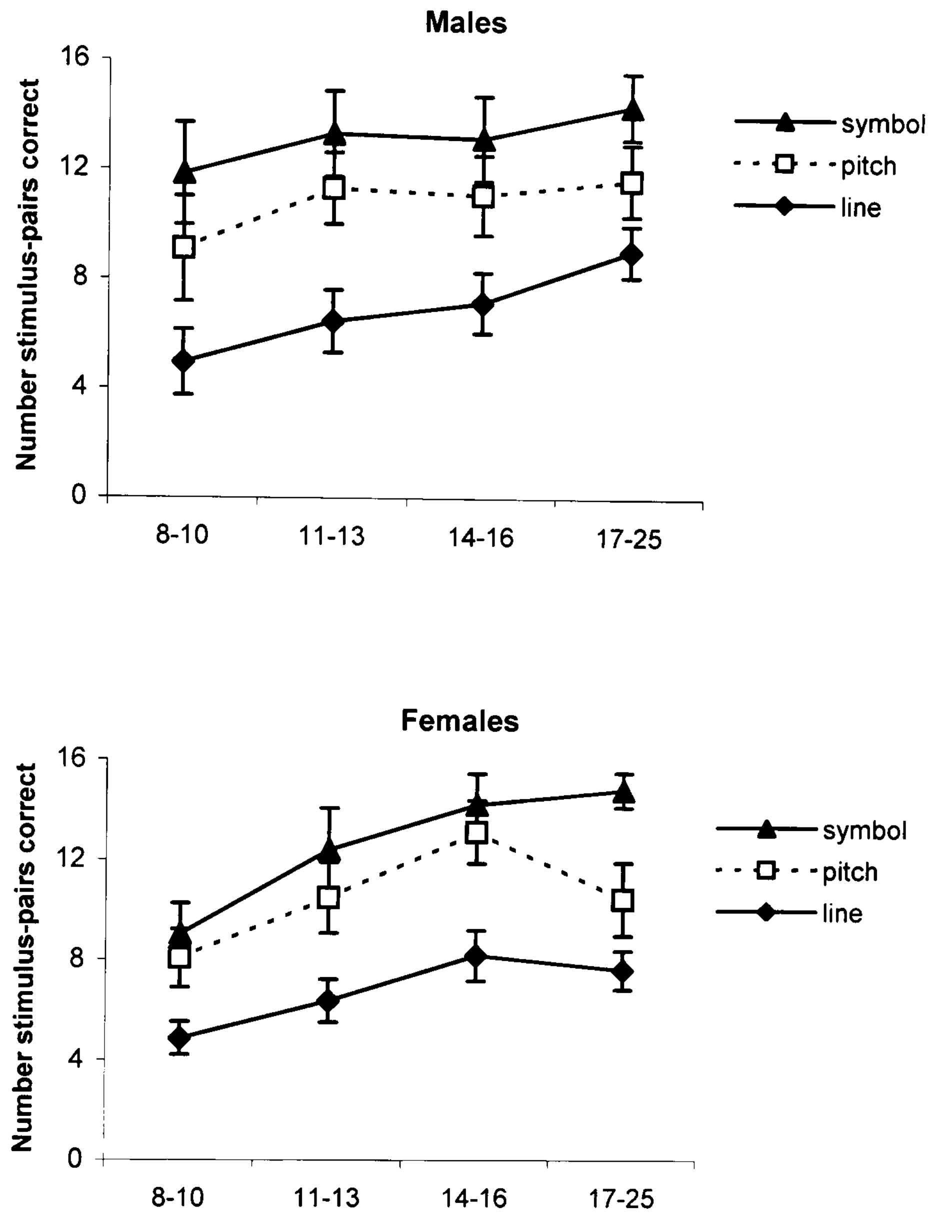


Figure 6.9 Mean number of correctly identified stimuli-animal pairs on the three associative learning tasks for male and females participants in each age group.

Between task comparisons

Comparisons between the three associative learning tasks for individuals were made using Wilcoxon's Signed Rank Tests. Within each age group, participants correctly identified more symbol-animal associations than line-animal pairs (z > 5.54, p < .0005), and pitch-animal pairs (z > 4.46, p < .0005). Furthermore, line-animal associations were more difficult than pitch-animal pairs (all z > 2.30, p < .02). This pattern was generally found for male and female participants in each age group, with the exception of comparisons between symbol-animal pairs and pitch-animal pairs, which did not significantly differ for females in the youngest three age groups (all z < 1.62, p > .10) and was marginal for males in the 14-16 year group (z = 1.93, p = .053).

Correlations

Bivariate correlations using the non-parametric Spearman's r_s were conducted to determine the relationship between performance on the associative learning tasks and age and IQ. Significant positive correlations were found between age and the number of correctly identified associations of the line angle, pitch, and symbol stimuli (all $r_s > .28$, $p < .0005$; males $r_s > .21$, $p < .03$; females $r_s > .33$, $p < .0005$). Significant positive correlations were also found between all IQ measures and performance on the three associative learning tasks (all $r_s > .29$, $p < .0005$; males $r_s > .30$, $p < .002$; females $r_s > .20$, $p < .03$). After controlling for the effects of age and FIQ, a significant positive correlation was found between the number of years of music training and Pitch Identification Task performance ($pr = .32$, males $pr = .25$, females $pr = .36$; all $p < .02$), but not Line Orientation or Symbol Task performance (all $pr < .02$, $p > .78$; males $pr < -.07$, $p > .21$; females $pr < .18$, $p > .07$).

All three tasks were found to correlate with each other in the TD sample ($r_s = .34$ to $.41$, all $p < .0005$). After covarying for age and FIQ, the association between Pitch Identification and Line Orientation ($pr = .19$, $p = .007$), and between Pitch Identification and Symbol ($pr = .17$, $p = .02$) still held, but not between Line Orientation and Symbol ($pr = .09$, $p = .18$). Associations between tasks were stronger in female participants (Pitch and Line, $pr = .23$, $p = .02$; Pitch and Symbol, $pr = .24$, $p = .01$; Line and Symbol, $pr = .19$, $p = .06$), than in male participants (Pitch and Line, $pr = .18$, $p = .09$; Pitch and Symbol, $pr = .10$, $p = .31$; Line and Symbol, $pr = .05$, $p = .66$), but not to a significant level (all $z_{r1\text{-}r2} < 0.99$, $p > .32$[5]).

6.4.6 *ASD and Control Results*

The Line Orientation and Pitch Identification Tasks were administered to all participants in the ASD and control groups. The Symbol Task was not administered to one 14-year-old boy from the ASD group and his performance on the two experimental tasks could not be directly compared to the control condition. Data from the Symbol Task were collected from all other participants.

The means and standard deviations for correct identification of the stimulus-animal pairs are shown in Table 6.8 for each group and for groups divided by high and low IQ. The accuracy data for the Line Orientation and Pitch Identification Tasks were normally distributed in both groups (Kolmogorov-Smirnov tests, p >.10). Data from the Symbol

---

5 The adjustment for comparisons between two independent partial correlations was made by adjusting the standard error to $1/\sqrt{(N\text{-}2\text{-}k)}$, where k is the number of partialled variables.

Task failed the tests of normality however ($p < .02$) and non-parametric statistical tests were therefore used throughout for consistency.

Table 6.8 Number of correctly identified items (maximum = 16)† on the three associative learning tasks for the ASD and control groups: Mean (*SD*)

	Group	N	Pitch[a]	Line	Symbol
All	ASD	31	9.71 (3.77)	6.00 (2.52)	9.87[1] (5.12)
	Control	31	10.10 (4.04)	6.29 (2.90)	11.35 (4.61)
High IQ	ASD	25	9.60 (3.71)	6.52 (2.52)	10.71[2] (5.03)
	Control	25	11.40 (3.21)	6.88 (2.86)	12.44 (4.39)
Low IQ	ASD	6	10.17 (4.36)	3.83 (0.75)	6.50 (4.32)
	Control	6	4.67 (2.07)	3.83 (1.47)	6.83 (2.14)

†Chance performance = 4. [1]N = 30, [2]N = 24
[a]$ASD_{high} < Control_{high}$, $p = .03$; $ASD_{low} > Control_{low}$, $p = .001$

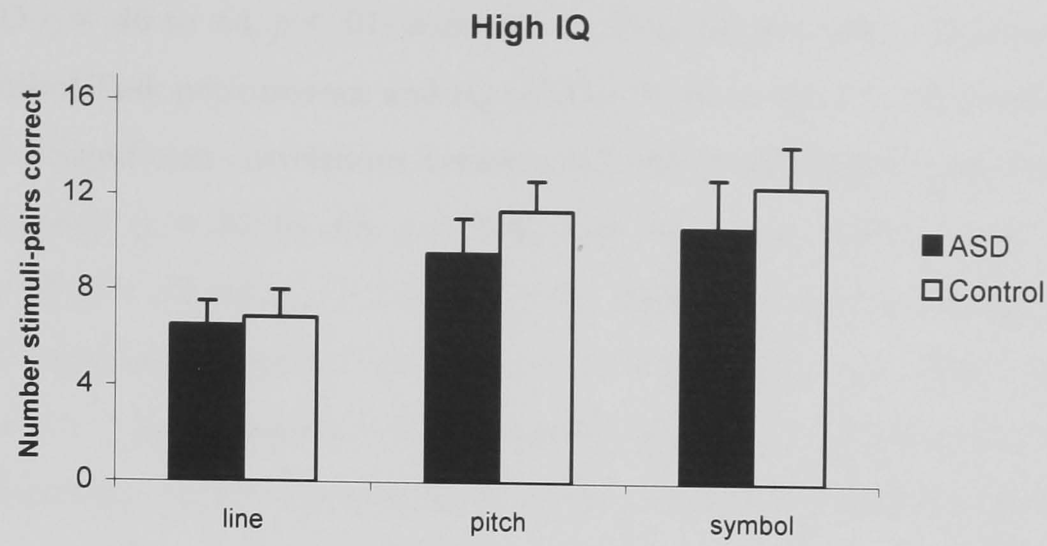
Binominal tests were conducted to determine the number of individuals performing above chance on each associative learning task ($p > .05$). The distribution of individuals performing above chance in the ASD and controls group was equivalent in the Line Orientation (36% versus 48%; $\chi^2 = 1.06, p = .15$, one-tailed) and Pitch Identification Tasks (81% versus 77%; $\chi^2 = 0.10, p = .38$, one-tailed). On the Symbol Task however, 60% of the ASD group performed above chance compared to 84% of the control group. The chi-square statistic confirmed that this distribution of participants differed significantly between groups ($\chi^2 = 4.32, p = .04$, two-tailed as no direction predicted).

No group differences were detected on the Line Orientation ($z = 0.49, p = .31$), Pitch Identification ($z = 0.56$, $p = .29$), or Symbol Task ($z = 1.25$, $p = .11$). When including participants with high IQ only, the ASD and control groups performed equivalently on the Line Orientation ($z = 0.66$, $p = .25$) and Symbol Task ($z = 1.17$, $p = .12$). Contrary to expectations, the high IQ control participants identified more pitch-animal pairs correctly on the Pitch Identification Task than ASD participants ($z = 1.87, p = .03$). The converse was found when comparing low IQ participants; the ASD group identified significantly more pitch-animal pairs than their low-functioning counterparts ($z = 2.35, p = .001$). This finding was only observed on the Pitch Identification Task; low IQ groups performed equivalently on the Line Orientation ($z = 0.17$, $p = .43$) and Symbol Tasks ($z = 0.89$, $p = .19$). The group difference on the Pitch Identification Task in the low IQ groups could not be explained by differences in musical experience as groups had received an equivalent number of years of musical training (see Appendix E).

As with the TD sample, a check was made whether participants were relying on relative processing to enable the correct identification of pitch- and line orientation-animal associations, rather than the intended absolute processing of stimuli. Removing the two extreme stimuli from the Line Orientation and Pitch Identification Tasks did not alter the overall findings between the ASD and control group (all $z < 1.08$, $p > .13$). High IQ control participants correctly identified significantly more pitch-animal pairs than high IQ ASD participants ($z = 1.77$, $p = .04$, one-tailed) and did not differ on line-animal pairs ($z = 0.57$, $p = .29$). The low IQ ASD participants now exceeded their matched control group on the number of line-animal pairs correctly identified ($z = 1.92$, $p = .03$, one-tailed), with a trend to identify more pitch-animal pairs ($z = 1.46$, $p = .07$). The findings suggest that the low-functioning participants in the ASD group may rely more on the absolute properties of line orientation, than the control group. As the pattern of results was similar for the two middle items as compared to all items, it can be proposed that the Line Orientation and Pitch Identification Tasks are assessing the ability to process the absolute properties of stimuli, and not merely relative processing.

Between task comparisons

Wilcoxon's Signed Rank Tests were used to compare individual performance between the three associative learning tasks. Participants in both groups identified significantly more symbol-animal pairs than line-animal pairs ($z > 3.69$, $p < .0005$). Learning a line angle association was significantly more difficult than musical note stimuli ($z > 3.69$, $p < .0005$). However, in contrast to the TD sample, neither group found pitch-animal pairs more difficult than symbol-animal pairs (ASD $z = 0.19$, $p = .85$; control $z = 1.41$, $p = .16$). The same pattern of results was found within high IQ participants, but differed when only including participants with low IQ (see Figure 6.10). Low IQ ASD participants identified more pitch-animal associations correctly than symbol-animal pairs ($z = 2.04$, N - ties = 5, $p = .04$), while the low IQ control participants found these two conditions equally difficult ($z = 1.6$, N - ties = 5, $p = .10$). The low IQ control group also found learning pitch-animal associations as difficult as line-animal pairs ($z = 0.53$, N-ties = 6, $p = .60$). Conversely, the low IQ ASD group correctly identified more pitch-animal pairs than line-animal pairs ($z = 2.20$, N - ties = 6, $p = .03$). Learning symbol-animal associations was easier than line-animal pairs in the low IQ control group ($z = 2.03$, N - ties = 5, $p = .03$). This was also apparent in the low IQ ASD group, but the difference between conditions did not reach significance ($z = 1.84$, N - ties = 4, $p = .07$).



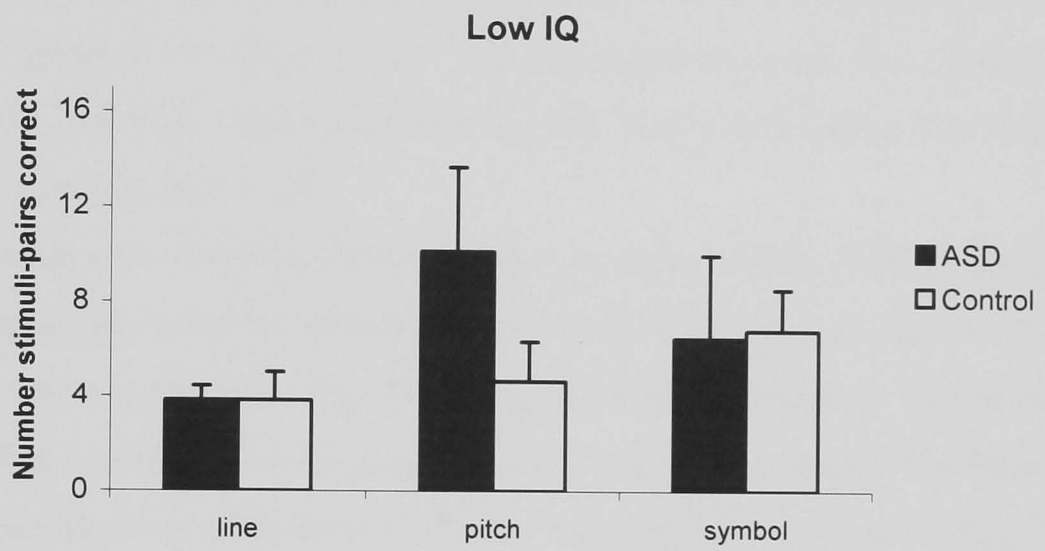


Figure 6.10 Mean number of correctly identified stimuli-animal pairings on the three associative learning tasks for the ASD and control groups divided by low IQ and high IQ. Error bars show the upper 95% confidence interval.

Correlations

No association was found between age and performance on each of the three associative learning tasks in the ASD group (r_s = -.16 to .08, all p > .40). A significant negative correlation was found between age and Symbol Task performance in the control group (r_s = -.38, p = .04), but no significant relationship was found with Line Orientation (r_s = -.11, p = .54) or Pitch Identification (r_s = -.33, p = .07).

Significant positive correlations were found between MA and Line Orientation Task performance in both groups (ASD r_s = .54, p = .002; control r_s = .51, p = .003). Positive correlations were observed between MA and Symbol Task performance, which reached significance for the ASD group only (ASD r_s = .43, p = .02; control r_s = .21, p = .26). No relationship was found between MA and Pitch Identification in the ASD group (r_s = -.04, p = .83), while a positive, but non-significant correlation was found between these two variables in the control group (r_s = .29, p = .12; z_{r1-r2} = 1.26, p = .20).

Strong positive correlations were found between Line Orientation and IQ in both groups (ASD $r_s = .46$ to $.64, p < .01$; control $r_s = .50$ to $.68, p < .006$). This was also found between Symbol Task performance and IQ (ASD $r_s = .46$ to $.62, p < .02$; control $r_s = .40$ to $.65, p < .03$). Significant correlations between IQ and Pitch Identification were found in the control group ($r_s = .52$ to $.65$, $p < .004$), but none were found in the ASD group ($r_s = -.18$ to $.09$, $p > .32$; all $z_{r1-r2} > 2.44$, $p < .01$). This is contrary to Heaton et al. (1998) who found significant positive correlations between FIQ and Pitch Identification performance ($r = .73$), and mostly notably, between PIQ and pitch performance ($r = .81$).

The relationship between performance on the Pitch Identification Task and years of formal music training was also explored. A significant positive correlation was found between these two variables in the control group ($r_s = .47, p = .008$), which was not evident in the ASD group ($r_s = -.06$, $p = .75$). No association between the number of years of music training and Line Orientation and Symbol Task performance was found in either group ($r_s = .01$ to $.20$, all $p > .27$).

Correlations were conducted between all three tasks. In the ASD group, performance on the Symbol Task correlated with the Line Orientation Task ($r_s = .52, p = .003$), while the association between Symbol and Pitch Identification approached significance ($r_s = .33$, $p = .08$). After considering covariance with FIQ and age, the association between Symbol and Pitch Identification performance became significant ($pr = .52$, $p = .005$), although the correlation between the Line Orientation and Symbol Tasks did remain significant ($pr = .27$, $p = .17$). No reliable correlation was found between the Pitch and Line Orientation Tasks in the ASD group ($r_s = .14$, $p = .46$; after controlling for age and FIQ, $pr = .16$, $p = .42$). In the control group, performance on the Symbol Task correlated significantly with the Line Orientation Task ($r_s = .39, p = .03$) and the Pitch Identification Task ($r_s = .40, p = .03$). No association was found between the Line Orientation and Pitch Identification Tasks ($r_s = .29, p = .11$). These significant correlations in the control group did not survive after covarying age and FIQ ($pr = -.16$ to $.15, p > .41$).

6.4.7 Summary and brief discussion of Pitch Identification and Line Orientation Tasks

It was predicted that a processing bias towards local details might enable formation of stable representations of individual notes, predisposing individuals to the acquisition of absolute pitch. It was considered that absolute line orientation might be a plausible analogue to absolute pitch in the visual modality, and both abilities were tested using a paired associative learning paradigm. A comparison task was included to obtain a measure of associative learning with stimuli that did not differ along a single dimension, and which would not benefit from weak coherence.

Performance on all three tasks showed developmental effects in the TD sample, with strong correlations between age and IQ and the ability to retain pitch-, line orientation-, and symbol-animal associations. The association between learning pitch-animal associations and developmental level was influenced somewhat by the musical experience of the individual, but the latter could not account for all the variance observed.

Overall, the ASD group did not show superior ability in retaining the absolute properties of pitch or line orientation, and were comparable in performance to their age- and IQ-matched controls. The small subgroup of low IQ participants with ASD did, however, show exceptional ability for their developmental level to retain pitch-animal associations. Performance on the Pitch Identification Task was therefore independent from IQ in the ASD sample, unlike the control group and TD sample. Somewhat surprisingly, the low IQ ASD participants performed at a higher level on the Pitch Identification Task than on the Symbol Task, which was designed to function as a control task, suggesting that music may be a privileged domain for this group.

Learning an association between a meaningless abstract form and an animal may have been relatively difficult for the ASD group for several reasons. First, although the possibility for semantic associations between pairs was minimised, it was likely that TD and control participants used this memory strategy to some extent. A specific difficulty has been found in ASD in their ability to create an organising structure to enable recall, notably of meaningless information (Ameli, Courchesne, Lincoln, Kaufman, & Grillon, 1988), which may have impacted on task performance. Secondly, as the abstract symbols were deliberately designed to share some local elements in order to reduce their discriminability, a local processing style may have hampered the ability to form symbol-animal associations.

Although the Line Orientation and Pitch Identification Tasks were designed to assess memory for the absolute properties of stimuli, it is plausible that individuals relied on the relative properties of some stimuli for accurate recall. Removing the two extreme items from each task did not alter the pattern of results, however, and suggests that the tasks did not depend wholly on relative processing of test stimuli. The low IQ ASD participants were found to identify significantly more line-animal pairs for middle items than their matched controls, providing some evidence for a local processing bias in the low-functioning subgroup of ASD.

For the majority of participants, absolute judgments of line orientation were more difficult than the absolute judgments of pitch. Within the scope of the task design, it was considered impossible to equate the magnitude of difference between lines of certain angles with difference in pitch. As the Line Orientation Task proved to be a difficult task, and did not distinguish between the ASD and control groups overall, this task was not taken forward as a key index of weak central coherence. Performance on the Pitch Identification Task (total number correct) was selected as the continuous variable, while performance of at least 75% pitch-animal pairs correctly identified (12 from 16) was used as the categorical variable as an indicator for weak central coherence.

6.5 Chord Sequence Task

It is well acknowledged in the study of music perception that melodies are typically experienced as gestalt phenomenon. Bever and Chiarello (1974) surmise "the fact that a melody is composed of a series of isolated notes is not relevant for naïve listeners; rather, they focus on the overall melodic contour" (p. 538). The Gestalt grouping principals of music, together with exposure to Western tonal music, lead to expectancies in how musical pieces should continue (Schmuckler, 1989). Schellenberg, Adachi, Purdy, and McKinnon (2002) point out that such expectations in music listening can also be influenced by the cognitive and perceptual disposition of individuals.

Research into harmonic priming has demonstrated that the processing of a target chord is facilitated (i.e., greater consonance and faster processing) when it follows another chord belonging to the same musical key (Tekman & Bharucha, 1992). Bigand and Pineau (1997) extended this finding to show that harmonic expectancies also occur at higher levels of musical structure. In their study, adult participants were asked to judge the consonance or dissonance of a final chord, while the global context (consisting of a six-chord sequence) was manipulated. It was found that the ability to attend to the local aspect of the task was greatly influenced by the global context. Even though the previous context was irrelevant to making consonant/dissonant judgments, target chords were more accurately and rapidly identified when they were harmonically related to the global context. The effect of priming was similar for musicians and non-musicians and was interpreted in terms of the internalised "harmonic regularities pervasive in Western musical pieces" (p. 1105).

Tillmann, Bigand, and Pineau (1998) investigated this result further by looking at the simultaneous effects of global and local contexts on harmonic priming. Adult participants made consonant/dissonant judgments of final target chords (the eighth chord of a sequence), while the preceding context was implicitly manipulated at both the global level (the initial sequence of six chords) and local level (the seventh chord). This rendered four different levels of congruity between the chord sequence and target chord, whereby the target chord was expected at both local and global levels, at neither the local nor global level, at the global level only, or at the local level only. Tillmann et al. found that harmonic expectancies were derived from both the global and local levels of musical structure. These two levels interacted with each other, such that the influence of the local context on processing a target chord was greater when the global context was unrelated to the target chord. The global context was influential in final chord processing, even when the local context was in conflict with the global context. The global context did not completely override the influence of the local context however, an effect that has also been found in discourse priming (Hess, Foss, & Carroll, 1995). Priming effects were also influenced somewhat by the tempo of the sequence: manipulations of the global context increased times to respond to consonant/dissonant judgments more when the sequences were presented at fast than slow tempo.

A theoretical explanation of the findings was provided by a connectionist model (Bharucha, 1987; Tillmann, Bharucha, & Bigand, 2000). In this model the implicit musical knowledge of the listener is represented as a network of interconnected units organized in three layers (i.e., tones, chords, keys). When a musical piece is presented, a spreading activation effect occurs whereby musical units that correspond to harmonically related chords are activated more strongly than units that correspond to unrelated chords. The activations are reported to accumulate but the influence of tempo on melodic expectancies suggests that the weightings of connections decay over time.

Music processing in ASD

As reviewed in Chapter 2 (Section 2.4.4), the study of local-global processing of musical stimuli in ASD has produced mixed findings: Mottron et al. (2000) found a local bias alongside preserved global processing; Foxton et al. (2003) found an absence of interference from the global level on the processing of local level information; and Heaton et al. (1999) found evidence of intact global processing. It is clear that further study is required to elucidate the influence of local and global contexts on the processing of musical pieces in ASD.

The musical priming paradigm, developed by Tillmann et al. (1998), raises interesting questions about whether individuals with autism would succumb to the same influences from global and local contexts on melodic expectancies. In an attempt to address these questions Heaton (2006) compared children with ASD and matched controls in their musical processing style using an adaptation of this paradigm. Participants were presented with sequences of eight chords and were asked to say whether the ending sounded right or wrong. Similar to Tillmann et al. the last target chord was either related or unrelated to the global context (the initial six chords), while simultaneously being related or unrelated to the local context (seventh chord). The task differed from Tillmann et al., as the consonant/dissonant judgment of the target chord was made explicitly in relation to the preceding context, rather than implicitly (i.e., a judgment about the final chord in isolation). Heaton found that both groups were influenced by local and global contexts in their judgments and no group differences were found. A significant interaction with tempo revealed stronger effects of the global context at slower presentation speeds. It was concluded that individuals with autism have preserved global processing within the music domain. The present study used the method adapted by Heaton in order to determine whether this finding of intact global processing is stable in a different sample of individuals with ASD and to explore individual differences.

Music processing in typical development

Studies have shown that melody perception, and the expectancies that are formed, change with development. In a review of the literature, Costa-Giomi (2003) reported significant improvements between the ages of 8 and 9 years in the ability to perceive harmony. With development, a shift in attention is made from perceiving concrete elements such as rhythm, pitch and contour, to more abstract elements such as harmony and key. This change has been attributed to the accumulative effects of exposure to Western tonal music and the underlying rules that structure this music (Sloboda, 1985). Costa-Giomi stated that greater memory capacity along with the ability to use strategies to direct attention to relevant aspects of the music, also contribute to this developmental change.

In a study of melodic expectancy, Schellenberg et al. (2002) compared groups of younger (7 to 8 years) and older (10 to 11 years) children and adults in their ability to rate how well individual tones continued short melodies. They found that participants of all ages expected the next tone in a melody to be close in pitch to the tone heard most recently, but this expectation was stronger for older than younger children and stronger for adults than the older children. With increases in age and exposure to music, expectancies became influenced by additional properties of the melodies. For example, expectations of reversals of pitch direction were strong in adults, but not in the two groups of children, who did not differ from one another.

Predictions

The present study examined the respective roles of global and local contexts and tempo on melodic expectancies in individuals with ASD or TD. Using the musical priming paradigm developed by Tillmann et al. (1998), it was predicted that a bias towards local processing (i.e., weak coherence) would result in relatively more congruent judgments in conditions where the local context was related to the target chord than when the local context was unrelated. Conversely, a bias towards global processing (i.e., strong coherence) would produce relatively more congruent judgments in conditions where the global context was related to the target chord then when the global context was unrelated. In line with the findings of Heaton (2006), the influence of the global context was predicted to be strongest at slow tempo, whereas the influence of the local context was expected to be greater at fast tempo.

As harmonic perception has been found to develop with age (Costa-Giomi, 2003; Schellenberg et al., 2002), the influence of the global context on melodic expectancies is predicted to increase with development in the TD sample. TD participants are expected to make more congruent judgments to target chords that are related to the preceding context at a global level than at the local level, with this response pattern being more prevalent in older than younger participants. Although studies of music processing in ASD have produced mixed findings, weak central coherence theory predicts that individuals with ASD would show relatively less influence from the global context on the processing of final chords and greater influence from local level information, compared to age- and IQ-matched control participants.

6.5.2 Method

Materials

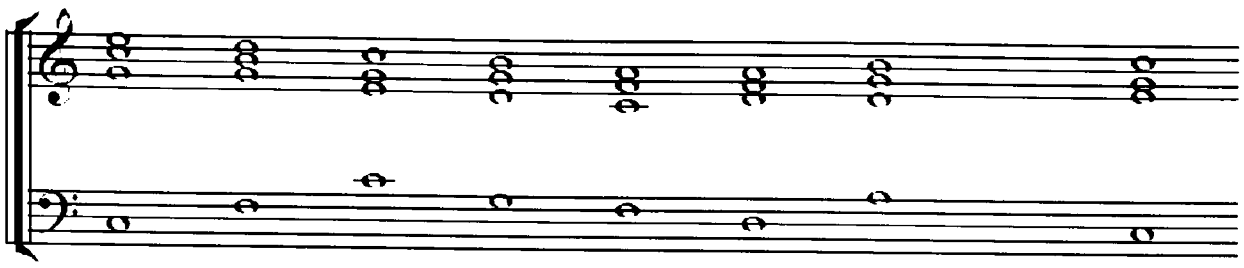
Stimuli consisted of 24 chord sequences composed of seven chords followed by a target chord. The stimuli were taken from Heaton (2006) following on from the work of Tillmann et al. (1998). Experimental conditions were made by systematically varying the local and global relatedness of the preceding harmonic context to the target chord. This resulted in four conditions in which the target chord was related to the context: (1) at both local and global levels (GRLR[6]); (2) at the global level only (GRLU); (3) at the local level only (GULR); or (4) at neither local nor global levels (GULU). See Figure 6.11 for examples of each experimental condition.

Chord sequences were recorded from a Casiotone 202 electronic keyboard (acoustic piano setting) onto a Sony Minidisk recorder. The individual chord sequences were played to participants through the Minidisk player using high quality speakers. To examine the influence of speed on music processing, the chord sequences were presented at either slow or moderate tempo (approximately 4 seconds and 7 seconds in total duration, respectively). The final set of stimuli therefore consisted of six exemplars of the four experimental conditions, three at *slow* presentation, and three at *fast* presentation. The chord sequences were administered to participants in a fixed random order.

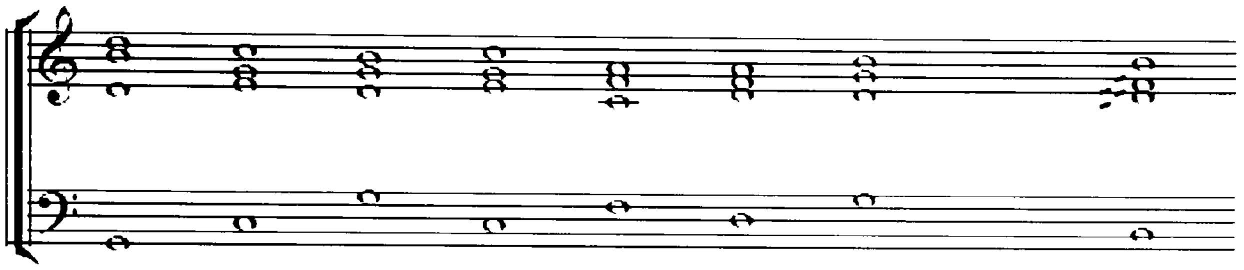
---

[6] G = globally, L = locally, R = related, U = unrelated.

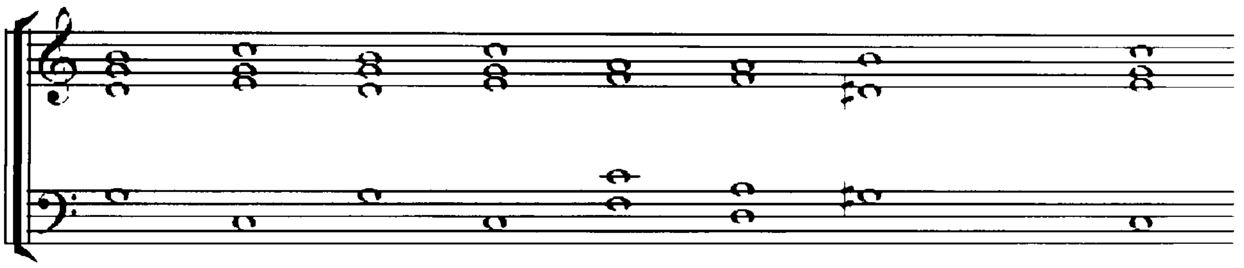
GRLR



GULU



GRLU



GULR

Figure 6.11 Examples of the four experimental conditions from the Chord Sequence Task.

Procedure

Participants were initially told that they would soon hear a short piece of music followed by a pause and then a final sound. Their task was to judge whether the final target chord sounded right or wrong. Participants were thus told, "*Listen to the music and tell me whether the ending sounds right*". The first trial was played and the experimenter ensured the participant understood the task requirements before continuing. To make the task as open ended as possible no feedback was provided. Furthermore, no time limit to respond was imposed but participants were encouraged to give their first impression of the correctness of the final sound. If requested, the sequence could be played again but only one repetition of each trial was allowed. The number of repetitions throughout the task was also limited to three occasions. The experimenter recorded participants' responses on individual score sheets.

GRLR and GULU trials were categorised as control conditions as the target chord was unambiguously correct or incorrect, respectively. Data collected on the open-ended conditions GRLU and GULR could only be judged reliable if it was established that participants could achieve a good level of performance on these control trials. Participants were therefore only included in the analysis if they were able to judge at least 4 of the 6 GRLR trials as correct, together with at least 4 of the 6 GULU trials as incorrect.

6.5.3 Typical Development Results

Nine participants from the TD sample were not administered the Chord Sequence Task due to time constraints. Of the 195 participants who received the task, 45 (23%) failed the inclusion criteria. These participants were younger ($M = 13.6$ years, $SD = 4.8$) than those who passed the inclusion criteria ($M = 15.0$ years, $SD = 4.3$), although the difference did not reach conventional levels of significance ($t_{(193)} = 1.91$, $p = .06$). The excluded participants had significantly lower FIQ ($M = 103.8$, $SD = 13.3$) and VIQ ($M = 105.2$, $SD = 16.0$) than participants who passed (FIQ $M = 109.0$, $SD = 14.5$, $t_{(193)} = 2.17$, $p = .03$; VIQ $M = 111.1$, $SD = 15.3$, $t_{(193)} = 2.23$, $p = .03$). Furthermore, the excluded participants had received significantly fewer years of music training ($M = 1.10$ years, $SD = 1.81$) than included participants ($M = 1.88$, $SD = 2.41$; $z = 2.36$, $p = .02$). Descriptives (and statistical comparisons) of participants who passed the inclusion criteria (N = 150) are presented in Appendix J, by age group and gender.

The mean number of target chords judged to sound correct, across the four experimental conditions for each age group is shown in Table 6.9 and Figure 6.12. As the data failed tests of normality (Kolmogorov-Smirnov tests, $p < .02$), non-parametric statistical analyses were performed.

A significant effect of age group was found in the GRLR condition ($\chi^2 = 10.95$, $p = .01$). Although this control condition was used to screen participants, the oldest age group still made more correct judgments than the two youngest age groups (all $z > 2.00$, $p < .05$). No age effects were found in the remaining three conditions (all $\chi^2 < 4.60$, $p > .19$). When dividing the sample by gender (see Table 6.10), a significant effect of age group was found in the GRLR condition for females ($\chi^2 = 10.84$, $p = .01$), but did not reach significance for males ($\chi^2 = 6.77$, $p = .08$). A further significant effect of age group was found in the GULU condition for male participants ($\chi^2 = 7.66, p = .05$). No age group effects were found for male or female participants on the two open-ended conditions, GRLU and GULR (all $\chi^2 < 4.39, p > .21$).

Table 6.9 Number of target chords judged to sound correct in each condition of the Chord Sequence Task for each age group (maximum = 6): Mean (*SD*).

Age group (years)	N	GRLR[a]	GRLU	GULR	GULU
8-10	34	5.15 (0.66)	4.47 (1.48)	3.65 (1.35)	0.74 (0.79)
11-13	32	5.28 (0.77)	4.72 (1.28)	3.72 (1.49)	0.66 (0.75)
14-16	38	5.39 (0.72)	4.92 (1.17)	4.00 (1.41)	0.45 (0.65)
17-25	46	5.63 (0.53)	4.87 (1.11)	4.07 (1.51)	0.46 (0.72)

[a]8-10, 11-13 < 17-25, $p < .04$

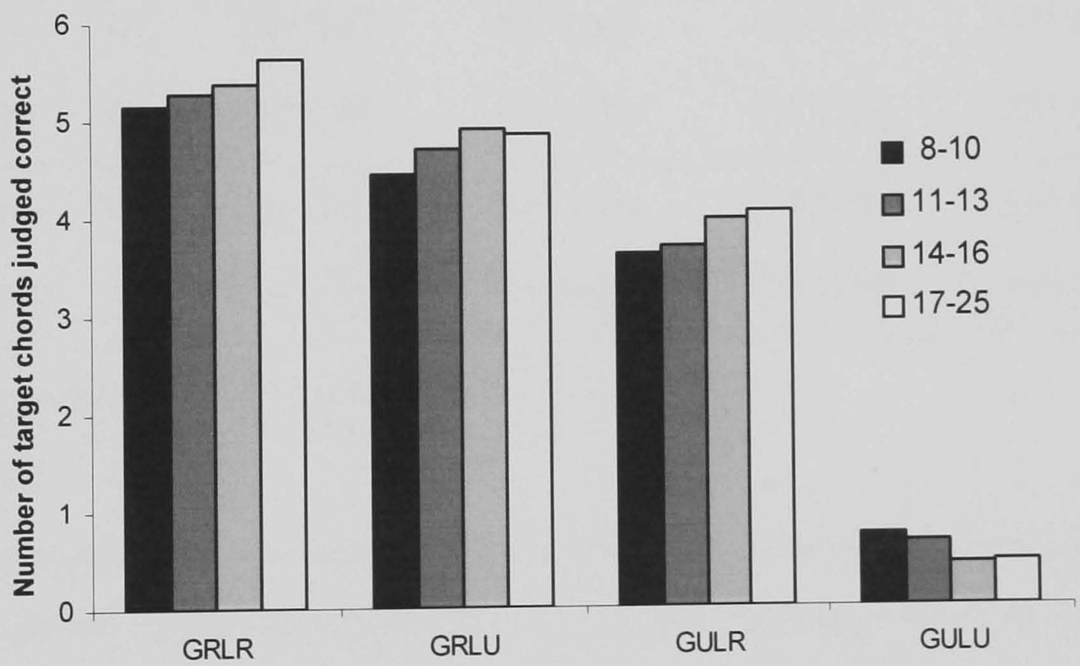


Figure 6.12 Mean number of target chords judged to be correct in each condition of the Chord Sequence Task for each age group (maximum= 6).

Male and female performance on the two open-ended conditions of the Chord Sequence Task was contrasted overall and within each age group (Table 6.10). A main effect of gender was found in the GRLU condition with females judging significantly more target chords as sounding correct than males ($z = 2.01$, $p = .04$). No significant main effects of gender were found in the other three conditions (all $z < 1.40, p > .15$). Females in the 14-16 year age group had significantly more years of formal music training than males (see Appendix J). No gender effects were found in this age group however, with only a trend showing females rated GRLU sequences as sounding correct more often than males ($z = 1.81, p = .07$). Male and female participants within each age group were found to perform equivalently in the GRLU and GULR conditions (all $z < 1.57, p > .11$).

Table 6.10 Number of target chords judged to sound correct in each condition of the Chord Sequence Task for each age group split by gender (maximum = 6): Mean (*SD*).

Age group (years)		N	GRLR[b]	GRLU	GULR	GULU[c]
8-10	*male*	13	5.15 (0.80)	4.15 (1.07)	3.38 (1.45)	0.85 (0.90)
	female	21	5.00 (0.87)	4.59 (1.68)	3.73 (1.32)	0.68 (0.72)
11-13	*male*	15	4.93 (0.80)	4.47 (1.51)	3.87 (1.30)	0.87 (0.64)
	female	17	5.59[a] (0.62)	4.94 (1.03)	3.59 (1.66)	0.47 (0.80)
14-16	*male*	23	5.30 (0.82)	4.70 (1.15)	3.74 (1.36)	0.57 (0.73)
	female	15	5.53 (0.52)	5.27 (1.16)	4.40 (1.45)	0.27 (0.46)
17-25	*male*	24	5.58 (0.50)	4.83 (1.17)	4.33 (1.40)	0.33 (0.64)
	female	22	5.68 (0.57)	4.91 (1.06)	3.77 (1.60)	0.59 (0.80)
All	*male*	75	5.29 (0.75)	4.60[a] (1.22)	3.89 (1.39)	0.60 (0.74)
	female	75	5.48 (0.60)	4.92 (1.27)	3.87 (1.50)	0.52 (0.72)

[a]Within (age) group effect of gender significant, $p < .05$
[b]Females: 8-10 < 11-13, 14-16, 17-25, $p < .05$
[c]Males: 17-25 < 11-13, $p < .007$

The four experimental conditions had the same effect on target chord judgments within all age groups. Wilcoxon's Signed Ranks Tests confirmed that for each age group, significantly more GRLR target chords were judged correct than GRLU target chords (all $z > 2.09, p < .04$). In turn, more GRLU target chords were judged correct than GULR target chords (all $z > 2.84, p < .005$). In addition, more GRLU and GULR target chords were judged correct than GULU (all $z > 4.80, p < .0005$). This pattern was found for male and female participants overall (all $z > 3.38, p < .001$), although exceptions occurred when analysing by gender and age group. Of note, no difference between GRLU and GULR conditions were found for males in the 11-13 and 17-25 age groups, and for females in the 8-10 and 14-16 age groups (all $z < 1.84, p > .06$).

Effect of presentation speed

Table 6.11 presents the mean number of target chords judged to sound correct for fast and slow presentation speeds, across the two open-ended conditions, overall and for each age group. Wilcoxon's Signed Ranks Tests were conducted to examine the effect of presentation speed within each condition. In the whole TD sample, GRLU target chords were more likely to be judged as sounding correct with slow than fast presentations ($z = 3.10$, $p = .002$). Conversely, GULR target chords were more likely to be judged as sounding correct with fast rather than slow presentations ($z = 2.38$, $p = .02$). The interaction between condition and presentation speed for the TD sample is presented in Figure 6.13. This pattern was found in female participants ($GRLU_{slow} > GRLU_{fast}$, $z = 2.55$, $p = .01$; $GULR_{fast} > GULR_{slow}$, $z = 2.99$, $p = .003$), but only in the GRLU condition for males ($GRLU_{slow} > GRLU_{fast}$, $z = 2.00$, $p = .05$). No effect of tempo was observed in the GULR condition for males ($z = 0.23$, $p = .82$).

Table 6.11 Target chords judged to sound correct in each open-ended condition and presentation speed in the Chord Sequence Task for each age group (maximum = 3): Mean (*SD*).

Age group (years)	Tempo	GRLU	GULR
8-10	*slow*	2.38 (0.78)	1.53[a] (0.99)
	fast	2.09 (0.97)	2.12 (0.77)
11-13	*slow*	2.44 (0.88)	1.78 (1.04)
	fast	2.28 (0.73)	1.94 (0.84)
14-16	*slow*	2.61 (0.64)	2.03 (0.79)
	fast	2.32 (0.84)	1.97 (0.91)
17-25	*slow*	2.59 (0.65)	1.93 (1.06)
	fast	2.28 (0.89)	2.13 (0.72)
All	*slow*	2.51[a] (0.73)	1.83[a] (0.99)
	fast	2.25 (0.86)	2.05 (0.81)

[a]Within (age) group effect of tempo significant, $p < .05$

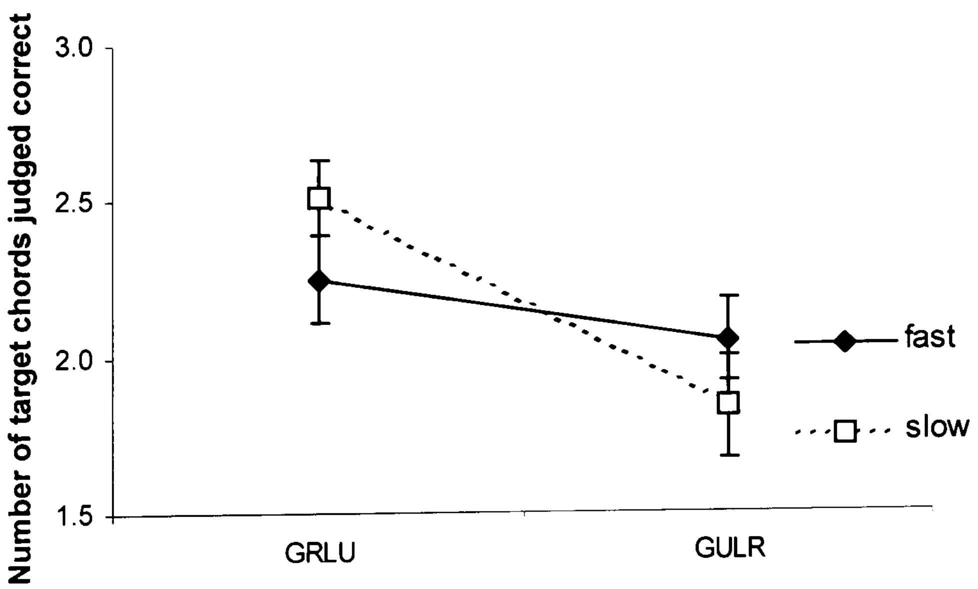


Figure 6.13 Mean number of target chords judged to sound correct for fast and slow presentation speeds, for GRLU and GULR chord sequence conditions (maximum = 3). Error bars show 95% confidence intervals.

The effect of presentation speed within condition was also examined for each age group. The finding that GRLU target chords were more likely to be judged as correct with slow than fast presentations approached significance for the two oldest age groups (14-16 years, $z = 1.85$, $p = .06$; 17-25 years, $z = 1.83$, $p = .07$) and the youngest age group ($z = 1.76$, $p = .08$). The finding that GULR target chords were judged correct more often at fast rather than slow presentations was only apparent in the youngest age group ($z = 2.74$, $p = .006$).

Correlations

Two independent measures of global and local processing bias were taken from the Chord Sequence Task for correlational analyses. The number of correct judgments made in the GRLU condition (fast and slow tempo combined) was used as an indicator of global preference, while the number of correct judgments made to the GULR condition (fast and slow tempo combined) was used as an indicator of local preference. While these measures are in principle independent, a significant positive correlation was found ($r_s = .28$, $p = .001$; males $r_s = .31$, $p = .008$; females $r_s = .26$, $p = .03$), suggesting an overall response bias towards saying yes or no to the compatibility of the target chord in the two open-ended conditions.

No significant correlation was found between age and the global score ($r_s = .12$, $p = .16$), or the local score ($r_s = .14$, $p = .09$). When analysing by gender however, significant positive correlations were found for males (global $r_s = .24$, $p = .04$; local $r_s = .26$, $p = .03$), but not for females (global $r_s = .02$, $p = .86$; local $r_s = .06$, $p = .64$). Fisher's z transformation showed the strength of the coefficients did not differ significantly between genders (all $z_{r1\text{-}r2} < 1.35$, $p > .18$).

A significant positive correlation was found between the global score and FIQ ($r_s = .17$, $p = .04$), but did not remain significant by gender (males $r_s = .22$, $p = .06$; females $r_s = .19$, $p = .13$). PIQ also showed a significant positive correlation with the global score ($r_s = .20$, $p = .01$; males $r_s = .24$, $p = .04$; not significant in females $r_s = .21$, $p = .08$). In contrast, VIQ showed no reliable correlation with the global score ($r_s = .09$, $p = .28$, males $r_s = .16$, $p = .16$, females $r_s = .04$, $p = .73$). Furthermore, when the Block Design subtest was removed from the FIQ score, no reliable correlation was observed with the global score ($r_s = .11$, $p = .20$; males $r_s = .16$, $p = .17$; females $r_s = .07$, $p = .55$). The local score did not relate to any IQ measure (all $r_s < .08$, $p > .35$; males $r_s < .13$, $p > .27$; females $r_s < .06$, $p > .62$).

It is known that music perception can be influenced by the degree of musical expertise of the individual (Bever & Chiarello, 1974). Females in the TD sample tended to have more years of music experience than males (2.2 years versus 1.6 years), with the difference approaching significance ($z = 1.79$, $p = .07$). The general finding of a greater bias towards global processing in females compared to males could therefore be attributed to differences in music training. It was observed that the correlation between music experience and the global score was significant in males ($r_s = .23$, $p = .05$) but not in females ($r_s = .18$, $p = .13$), although the strength of the correlation coefficients did not differ significantly ($z_{r1-r2} = 1.50$, $p = .13$).

6.5.4 *ASD and Control Results*

The Chord Sequence Task was administered to 30 participants in the ASD group and 30 participants in the control group. Eight participants from each group (27%) failed the inclusion criteria of judging at least 4 of the 6 GRLR trials as sounding correct and 4 of the 6 GULU trials as sounding incorrect. The excluded participants from the ASD group were younger ($M = 13.4$ years, $SD = 2.7$) than those who had passed the inclusion criteria ($M = 15.3$ years, $SD = 2.2$), although the difference did not quite reach conventional levels of statistical significance ($t_{(28)} = 1.98$, $p = .06$). The included and excluded participants from the ASD group were equivalent on all IQ measures (all $t_{(28)} < 0.30$, $p > .77$) and had received a similar number of years of music training ($z = 1.12$, $p = .26$). Similarly, the included and excluded participants from the control group were equivalent in age and IQ (all $t_{(28)} < 0.64$, $p > .53$), and had received the same number of years of music training ($z = 0.34$, $p = .74$).

ASD and control groups comprising participants who passed the inclusion criteria did not differ in age, IQ, or mean years of musical training (see Appendix J). The mean number of target chords judged to sound correct in each of the four conditions of the Chord Sequence Task is presented in Table 6.12 and Figure 6.14. As the data failed tests of normality (Kolmogorov-Smirnov tests, ASD $p < .02$, control $p < .10$), non-parametric tests were used.

Opposite to predictions of weak central coherence theory, the ASD group judged more target chords to sound correct in the GRLU condition than the control group ($z = 1.78$, $p = .04$, one-tailed). No group differences were found on the remaining conditions (all $z < 1.44$, $p > .08$). The data were also examined for high- and low-functioning participants within each group (see Table 6.12). No group differences were detected within participants with high IQ (all $z < 1.24$, $p > .11$). Surprisingly, the low-functioning ASD group reported more target chords in the GRLU condition to sound correct than the control group. Wilcoxon's Rank-Sum Test for small sample sizes (Howell, 2002, p. 710) found this difference to be statistically significant at the one-tailed level ($z = 1.73$, $p = .04$). There was also a trend for the low IQ ASD group to judge more GRLR target chords as sounding correct than the control group ($z = 1.56$, $p = .06$), while no differences were observed on the GULR condition ($z = 0.29$, $p = .39$) or GULU condition ($z = 0.68$, $p = .25$).

Table 6.12 Number of target chords judged to sound correct in each condition of the Chord Sequence Task for each group (maximum = 6): Mean (*SD*).

	Group	N	GRLR	GRLU[a]	GULR	GULU
All	ASD	22	5.45 (0.80)	5.00 (1.07)	3.82 (1.18)	0.50 (0.74)
	Control	22	5.27 (0.77)	4.45 (1.01)	3.55 (1.44)	0.82 (0.80)
High IQ	ASD	18	5.39 (0.85)	4.94 (1.11)	3.89 (1.08)	0.50 (0.71)
	Control	18	5.39 (0.70)	4.61 (0.98)	3.56 (1.42)	0.78 (0.73)
Low IQ	ASD	4	5.75 (0.50)	5.25 (0.96)	3.50 (1.73)	0.50 (1.00)
	Control	4	4.75 (0.96)	3.75 (0.96)	3.50 (1.73)	1.00 (1.15)

[a]ASD > Control, $p = .04$; ASD_{low} > $Control_{low}$, $p = .04$

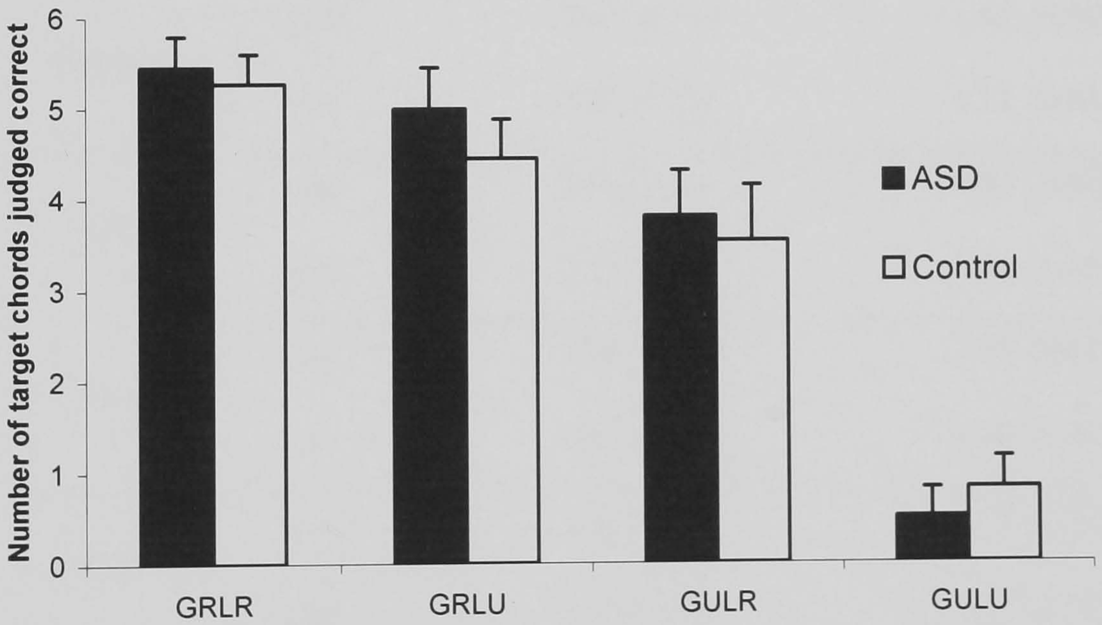


Figure 6.14 Mean number of target chords judged to sound correct for each condition of the Chord Sequence Task for each group (maximum = 6). Error bars show the upper 95% confidence interval.

Wilcoxon's Signed Ranks Tests were conducted within each group to examine the relative effect of each condition. In the control group significantly more GRLR target chords were judged correct than GRLU target chords ($z = 2.97$, $p = .003$). In the ASD group however, these two conditions produced an equal number of correct judgments ($z = 1.53$, $p = .13$). Both groups judged significantly more GRLU target chords as sounding correct than GULR target chords (ASD $z = 3.02$, $p = .003$; control $z = 2.17$, $p = .03$). Furthermore, GULR chords were judged more often as sounding correct than GULU chords for both groups (ASD $z = 4.11$, control $z = 4.03$, $p < .0005$).

Effect of presentation speed

Table 6.13 presents the mean number of target chords judged to sound correct for fast and slow presentation speeds, in the two open-ended experimental conditions, for each group. No difference between fast and slow presentations was found in either group for the GULR condition (Wilcoxon's Signed Ranks Tests, ASD $z = 0.15$, $p = .89$; control $z = 1.28$, $p = .20$), or the GRLU condition (ASD $z = 1.00$, $p = .32$; control $z = 0.98$, $p = .33$). No effect of presentation speed in the two conditions was also found within high-functioning participants (ASD all $z < 1.30$, $p > .19$; control $z < 1.32$, $p > .18$), nor low-functioning participants (ASD all $z < 1.10$, $p > .30$; control $z < 1.10$, $p > .30$).

Table 6.13 Number of target chords judged to sound correct in each condition and presentation speed in the Chord Sequence Task for each group (maximum = 3): Mean (*SD*).

	Group	Tempo	GRLU	GULR
All	ASD	*slow*	2.55 (0.74)	1.86 (0.94)
		fast	2.45 (0.60)	1.95 (1.05)
	Control	*slow*	2.27 (0.77)	1.82 (0.85)
		fast	2.18 (0.80)	1.73 (0.88)
High IQ	ASD	*slow*	2.56 (0.78)	1.89 (0.90)
		fast	2.39 (0.61)	2.00 (1.08)
	Control	*slow*	2.39 (0.78)	1.78 (0.81)
		fast	2.22 (0.81)	1.78 (0.81)
Low IQ	ASD	*slow*	2.50 (0.58)	1.75 (1.26)
		fast	2.75 (0.50)	1.75 (0.96)
	Control	*slow*	1.75 (0.50)	2.00 (1.15)
		fast	2.00 (0.82)	1.50 (1.29)

Correlations

The mean number of correct judgments made to target chords presented in the GRLU and GULR conditions (fast and slow tempo combined) was taken as an indication of global and local processing bias respectively. While a positive but non-significant correlation was observed in the ASD group between these two measures of local and global processing ($r_s = .24$, $p = .28$), no relation between these measures was found in the control group ($r_s = .06$, $p = .79$; although not differing in strength, $z_{r1\text{-}r2} = 0.77$, $p = .44$).

The correlation coefficients between age and the global score were of similar magnitude to that in male TD participants, but did not reach significance in the ASD group ($r_s = .33$, $p = .14$) or the control group ($r_s = .26$, $p = .24$). A positive correlation was found between MA and the global score for control participants ($r_s = .53$, $p = .01$), but was not observed in ASD participants ($r_s = -.16$, $p = .47$). Fisher's z transformation showed the correlation coefficients differed significantly in strength ($z_{r1\text{-}r2} = 2.32$, $p = .02$). Moderate but non-significant correlations were found in the control group between the local score and age ($r_s = .29$, $p = .20$) and MA ($r_s = .36$, $p = .36$), which were not found in the ASD group (all $r_s < .13$, $p > .47$; although did not differ significantly in strength, all $z_{r1\text{-}r2} < 0.77$, $p > .44$).

Negative correlations were observed between the IQ measures and the global score in the ASD group, although none was found to reach significance ($r_s = -.07$ to $-.32$, all $p > .14$). In contrast, positive correlations were found in the control group between the global score and IQ (FIQ $r_s = .39$, $p = .08$; FIQ-BD $r_s = .24$, $p = .28$; PIQ $r_s = .43$, $p = .05$; VIQ $r_s = .27$, $p = .23$; all $z_{r1\text{-}r2} < 1.87$, $p > .05$). No significant correlations were found between the IQ measures and the local score in the ASD group ($r_s = -.06$ to $.22$, all $p > .31$), nor in the control group ($r_s = -.09$ to $.06$, all $p > .67$). Furthermore, the mean number of years of music training did not correlate with the global or local score in either group (ASD $r_s < .19$, $p > .39$; control $r_s < .10$, $p > .66$).

6.5.5 Summary of Chord Sequence Task

The predicted developmental effects on the Chord Sequence Task were not strongly supported by the data. In the TD sample, age effects were only found in control conditions, where the global and local contexts were both either concordant or discordant to the target chord (GRLR, GULU). In the two open-ended conditions, in which only the global or local level was concordant to the target chord (GRLU, GULR), no overall effects of age were found. When analysing by gender however, age effects were confirmed in male participants, but not females. Older males showed a tendency to make more "correct" judgments when the target chord was compatible at either the global or local level than younger males. A gender effect was found, with females showing a greater tendency than males to judge correct those target chords that were consistent only at the global level (GRLU).

The tendency to judge target chords that were compatible only at the global level (GRLU) as correct showed a positive correlation with IQ in the TD sample, in particular with PIQ and the Block Design subtest. This finding was mirrored in the control group, but not in the ASD group where negative and non-significant correlations were found between IQ and the tendency to make global judgments. The tendency to accept target chords that were compatible only at the local level (GULR) showed no relationship with IQ across all participants. It may be that integration of the seven chords takes central processing resources in TD and control participants. Judgments made to the local level place no such demands, hence no IQ effects were found.

In the TD sample the predicted effects of tempo on local and global processing were confirmed. The influence of the global context was stronger for slow compared to fast presented sequences. Conversely, the influence of local context was stronger for fast compared to slow presented sequences. Possibly the fast tempo was too fast for some participants to process and integrate the global form completely, allowing local context to predominate. Within-group effects of tempo on global and local processing were not found in the ASD or control group.

The important finding that emerged from the Chord Sequence Task was that both global and local contexts had an influence on musical expectancies. Significantly more target chords were judged as sounding correct when both levels were concordant, versus only the global level being concordant (GRLR > GRLU). Similarly, more target chords were judged as correct when only the local context was concordant, compared to when neither context levels were concordant (GULR > GULU). The former effect was larger than the latter, hence it seems that global coherence in music processing had greater prominence than local coherence. This pattern of findings was found not only in the TD group, but also in the ASD and control groups, and suggests that the influence of global level information in music processing was pervasive across groups. Moreover, individuals with ASD, and in particular those with low IQ, were influenced by the global context to a greater extent than the control group. Significantly more target chords were judged as sounding correct when they were concordant at the global level only (GRLU) in the ASD group relative to the control group. Although contrary to predictions of weak central coherence theory, this finding replicates Heaton (2006) and adds to the evidence of intact global processing of music in ASD.

Key dependent variables taken from the Chord Sequence Task for subsequent analyses were the total number of target chords judged correct when compatible at the global level only (GRLU), and when compatible at the local level only (GULR). Two independent measures of the degree of processing bias towards the global and local levels were therefore obtained for each individual. Individuals were categorised as showing weak central coherence if they displayed a bias towards local processing on either of the two open-ended conditions. That is, fewer than four from six affirmative responses to target chords presented in the GRLU condition, or more than four from six affirmative responses to target chords presented in the GULR condition.

6.6 General Discussion

A summary of experimental findings from the four low-level auditory coherence tasks can be found in Table 6.14 for the TD sample and in Table 6.15 for the ASD and control groups (see Appendix Q for effect sizes between ASD and control groups). Table 6.16 presents the percentage of participants in each group categorised as showing weak central coherence on each measure.

Three auditory tasks required the ability to focus on local elements and disregard context, to either isolate a sound from its presented configuration (i.e., a non-word or chord), or to process the absolute quality of a sound. Performance on these tasks was strongly associated with age and IQ in the TD sample, with greater discrimination accuracy and segmentation skills found in older and brighter individuals. However removing the effects of age, IQ, and music training still left a substantial proportion of the variance unexplained (see Table 6.14) which may reflect individual differences in cognitive style. No effects of gender were found on these measures. On an open-ended task of music processing, where participants judged whether the endings of chord sequences sounded correct, developmental effects were less obvious. Females were influenced by the global context to a greater extent than males, while a relationship between age/IQ and global processing was found in males but not females.

In general, individuals with ASD did not show superior disembedding of auditory stimuli as predicted by weak central coherence theory. Control participants were somewhat better than ASD participants at identifying the presence of a phoneme within a non-word, while the two groups did not differ in the ability to determine the presence of a note within a chord. TD and control participants took longer to determine the presence of a phoneme when it was highly embedded as the medial or final sound, than compared to when the phoneme was the initial sound. Interestingly, this effect of position on time to disembed the target phoneme was absent in individuals with ASD. This finding supports the prediction from weak central coherence theory, in that individuals with ASD are not hampered by greater embedding within the gestalt form. However, in the absence of superior performance overall on this measure, it is difficult to conclude whether this finding shows differences in perceptual style.

The small subgroup of low IQ individuals with ASD demonstrated superior ability to retain the absolute property of pitch information. This same group of individuals showed a bias towards global expectancies on the processing of chord sequences, providing evidence that local processing might coexist with intact global processing in autism, at least in processing musical stimuli.

The Chord Segmentation and Phoneme Segmentation Tasks were designed to be auditory analogues of the EFT; these tasks require the ability to disembed a local element from its surround. The tasks differ in how the local and global elements are presented, however. In both the Phoneme Segmentation and the EFT, the local element is known before the global form is presented. In contrast, in the Chord Segmentation Task the presentation of the global form (the chord) precedes that of the local part (the note). Discrimination accuracy for the presence of single notes within chords was low across all participants. This may relate to the presentation order of stimuli although further study is needed to confirm this. Heaton (2003) highlighted the importance of prior familiarisation of notes: children with autism showed superior ability to segment musical chords only when they were familiar with the target notes. Future studies should therefore take this into consideration, and provide participants with the note before the chord.

To conclude, the predictions from weak central coherence theory in the processing of low-level auditory information were largely not confirmed in the sample of individuals with ASD. Instead, processing tended to be captured by the global structure of sequentially presented musical stimuli, with little indication of superior segmentation abilities in ASD. Individual differences were found, that effects of age, IQ, and music training did not account for, and which might reflect differences in cognitive style. In Chapter 9 the inter-relations of the four low-level auditory coherence measures will be explored, together with investigations of task performance across modality (low-level visuo-spatial measures) and level (high-level visuo-spatial and verbal measures).

Table 6.14 Summary of low-level auditory task findings in the TD sample.

Auditory Task	Dependent Variable (Direction of high values)	Age effects?	IQ effects?	Gender effects?	Music training effects?	Proportion of variance not explained by age, IQ, gender (and music training)	Interactions
Phoneme Segmentation	A′ (Weak CC)	Yes, positive	Yes, positive	No	N/A	91%	Positive correlation with age in F only
	Relative effect of position (Strong CC)	No	No	No	N/A	99%	Negative correlation with FIQ in F only (trend)
Chord Segmentation	A′ (Weak CC)	Yes, positive	Yes, positive	No	Yes, positive	88%	
	Correct detections (Weak CC)	Yes, positive	Yes, positive	No	Yes, positive	88%	Positive correlation with age & IQ in M only
Pitch Identification	Total Correct (Weak CC)	Yes, positive	Yes, positive	No	Yes, positive	73%	Trend for F > M in 14-16 but explained by more years of music training
Chord Sequence[a]	Number GRLU judged correct (Strong CC)	No	Yes, positive	Yes, F > M	Positive in M only	94%	Positive correlation with age & PIQ in M only
	Number GULR judged correct (Weak CC)	No	No	No	No	99%	Positive correlation with age in M only

[a]N = 150 who pass inclusion criteria. Note: M = males, F = females

Table 6.15 Summary of low-level auditory domain findings in the ASD and Control groups.

Auditory Task	Dependent Variable (Direction of high values)	Group effects?	Age effects?	MA effects?	IQ effects?	Other comments
Phoneme Segmentation	A′ (Weak CC)	ASD < Controls (not as predicted)	No	Yes, positive	Yes, positive	$\text{ASD}_{\text{high}} < \text{Control}_{\text{high}}$ (not as predicted)
	Relative effect of position (Strong CC)	ASD < Controls	No	No	No	
Chord Segmentation	A′ (Weak CC)	No	No	No	No	
	Correct detections (Weak CC)	No	No	No	Positive in Controls only	
Pitch Identification	Total Correct (Weak CC)	No	No	No	Positive in Controls only	$\text{ASD}_{\text{low}} > \text{Control}_{\text{low}}$ $\text{ASD}_{\text{high}} < \text{Control}_{\text{high}}$ (not as predicted)
Chord Sequence[a]	Number GRLU judged correct (Strong CC)	ASD > Control (not as predicted)	No	Positive in Controls only	Positive in Controls only	$\text{ASD}_{\text{low}} > \text{Control}_{\text{low}}$ (not as predicted)
	Number GULR judged correct (Weak CC)	No	No	No	No	No

[a]N = 44, who pass inclusion criteria.

Table 6.16 Percentage (and N/total N) of participants in each group who showed weak central coherence on each task.

Auditory Task	Marker for weak CC	8-10	11-13	14-16	17-25	ASD	Control
Phoneme Segmentation	At least 12 from 15 phonemes in medial/final positions correctly disembedded	44% (24/54)	33% (14/43)	26% (13/51)	46% (26/56)	23% (7/30)	27% (8/30)
Chord Segmentation	At least 6 from 8 notes correctly disembedded	39% (21/54)	47% (20/43)	51% (26/51)	45% (25/56)	37% (11/30)	48% (15/31)
Pitch Identification[a]	At least 12 from 16 correctly identified	19% (10/54)	44% (19/43)	63% (32/51)	60% (33/55)	29% (9/31)	39% (12/31)
Chord Sequence	Fewer than 4 GRLU or more than 4 GULR target chords judged correct	50% (17/34)	56% (18/32)	50% (19/38)	48% (22/46)	46% (10/22)	41% (9/22)

[a]8-10 < 11-13, 14-16, 17-25, $p < .03$

Chapter 7. High-level visuo-spatial coherence tasks

7.1 Introduction

Chapter 7 presents experimental findings from three tasks designed to assess individual differences in the processing of meaningful visual information. A balance is achieved by selecting tasks that tap both local and global processing style. The first task examines differences in drawing style; that is whether an individual draws in a coherent, organised manner, versus a fragmented, piecemeal approach. Two different drawing stimuli are contrasted in order to examine the effect of meaning on drawing accuracy in recall. Participants are compared in the second task in their style of describing a detailed picture of a complex scene. The relative frequency of detailed descriptions is investigated, as well as whether an overall, global description of the picture was made. Memory for exact details within the picture, suggested to reflect local processing bias, is also examined. The final task requires the ability to cohere fragments of a line drawing in order to make sense of ambiguous visual information; a skill that is suggested to place demands on integrative global processing.

7.2 Drawing Task: Rey and Pram Figures

Drawing and copying tasks have been widely used to explore a variety of cognitive processes, including planning and organisational skills, problem-solving strategies, as well as perceptual, motor, and memory functions (for a review of cognitive models of drawing abilities, see Guerin, Ska, & Belleville, 1999). One such task that has had extensive use since its inception is the Rey-Osterrieth Complex Figure (Rey, 1941, cited in Lezak, 1995). The Rey figure was originally designed to assess perceptual organisational and visuospatial memory of brain-injured adults (Lezak, 1995; Spreen & Strauss, 1998). The figure can be described as an abstract geometric design that is structured around a base rectangle (see Figure 7.1). In the standard administration participants are required to copy the figure while it is in view, and then, without prior warning, reproduce it from memory (immediately and/or following a delay of 20-30 minutes). As well as assessing accuracy of the copy and recall, the quality of the construction can be evaluated.

The Rey figure can be perceived as a hierarchically organised structure consisting of global structures (e.g., the base rectangle) and local elements (e.g., single line details). Akshoomoff and Stiles (1995a) point out that the figure can be considered in terms of multiple hierarchical levels, but adults and children may differ in how they conceptualise which elements comprise each level. With respect to this organisation, various strategies have been identified in how individuals approach the figure. A global approach, for example, is shown by a tendency to draw the main structural elements first with the addition of local elements placed in relation to this framework. Conversely, individuals may copy the figure in a fragmented, part-orientated manner, initially drawing local details and failing to maintain their overall spatial organisation. An extreme piecemeal approach is often seen in very young children or individuals with right hemisphere damage (Akshoomoff, Feroleto, Doyle, & Stiles, 2002), although healthy adults still differ in the degree to which they use a local or global strategy. The method of construction in drawing the Rey figure was therefore taken in the present study as an indication of processing style in the visuo-spatial domain.

Drawing Style in Typical Development

Several studies have documented the normal developmental changes in children's ability to copy and recall the Rey figure (e.g., Akshoomoff & Stiles, 1995a, 1995b; Demsky, Carone, Burns, & Sellers, 2000; Karapetsas & Vlachos, 1997; Waber & Holmes, 1985, 1986). Typically, young children copy the Rey figure in a piecemeal fashion and appreciation of the configural form increases with age (Spreen & Strauss, 1998). Osterrieth (1944, cited in Lezak, 1995) was the first to obtain normative data for children aged four years up to adulthood. Children younger than seven years were found to draw discrete parts of the figure, without any notion of organisation. Older children would impose some structure in their drawing, with more than half of the children aged 13 years and upward starting by drawing the base rectangle and then adding details in relation to it.

Waber and Holmes (1985) undertook a normative study of children (N = 454, 5 to 14 years) to document aged-related changes in drawing the Rey figure, particularly in goodness of organisation and copying style. They found that by age 9 years, children could reliably reproduce all parts of the figure. Differences in performance after this age were determined by the effective use of organisation and planning strategies. Children who produced better copies were seen to view the base rectangle as a "salient organisational unit" and more often used continuous lines to draw the main structural elements of the figure. Waber and Holmes (1986) also described the development of children's ability to reproduce the Rey figure from memory. Typically all children from the age of 9 onwards recalled the main structures of the figure including the base rectangle. Incidental features were recalled less well and were more often drawn inaccurately or misplaced. Children were more likely to take a configural approach when drawing the figure from memory compared to their copy productions, although this shift in strategy was not so evident amongst the youngest children. Taking a piecemeal, or part-oriented approach when drawing the figure from memory was rare after the age of 9 years.

In a study of 160 children between 6 and 12 years of age, Akshoomoff and Stiles (1995a) attempted to describe more closely the processes used by children in drawing the Rey figure. They found that not only did accuracy improve with age, but also how children formed parts of the figure, and how the parts were then integrated, changed with development. Children as young as 6 years were found to reproduce nearly all of the elements of the figure and showed some ability to integrate the parts in their drawings. The most common strategy taken by children in the sample was different to the global approach shown in adults. In preference to drawing the entire base rectangle, the majority of children divided the figure into two or three subunits, and would complete one subunit before drawing the next. The authors argued that although this approach is less configural, it should not be interpreted as purely part-oriented but evidence of a different way of perceiving the figure.

Akshoomoff and Stiles (1995b) also examined the consistency of the strategy used by children in their sample across the copy and immediate recall trials. The way in which children started to draw the figure was generally consistent across memory and copy, although most children were more likely to use continuous lines when drawing the main structure of the figure from memory than direct copy, thus showing greater appreciation of the base rectangle. The method taken to copy the figure also influenced the way it was drawn from memory; children who took a very piecemeal and disorganised approach at copy were found to show poor immediate recall. An age effect was apparent with younger children recalling less of the figure than older children. Furthermore, across all children, configural elements were more likely to be remembered than local elements.

In addition to age, intellectual level is reported to contribute to performance on the Rey figure (Spreen & Strauss, 1998). Poulton and Moffitt (1995) reported a strong correlation between accuracy on the copy and recall trials and PIQ from the Wechsler scales. This correlation was most notable for the Block Design and Object Assembly subtests, suggesting the involvement of visuo-spatial and constructional abilities.

The role of gender on performance of the Rey figure is unclear. In their normative sample of 740 children aged 13 years, Poulton and Moffitt (1995) found males significantly outperformed females in recalling the Rey figure although the difference was marginal and not considered clinically significant. In a study of 840 children, aged 5 to 12 years, Karapetsas and Vlachos (1997) reported a gender difference at younger ages (6.5 to 9.5 years) with girls performing significantly better than boys. No measure of IQ was taken, however, which would have allowed for examination of FIQ/PIQ effects as a possible confound for sex differences. Demsky et al. (2000) also collected normative data on the

Rey figure for 432 children, aged 6 to 11 years. They found a steady increase in performance with age over their whole sample, but no gender differences were found.

Drawing Style in Autism

Several claims have been made in the literature that the drawing style of individuals with autism is unusual and, at times, precocious. A small proportion of individuals with autism have demonstrated exceptional drawing ability, out of line with their general intelligence (O'Connor & Hermelin, 1988; Sacks, 1995; Selfe, 1977). The drawings produced by these individuals are often visually realistic, which motivated Charman and Baron-Cohen (1993) to investigate whether the normal shift in drawing development from intellectual realism to visual realism occurred earlier in autism than in TD. They found no difference, however, in the production of intellectually and visually realistic drawings made by children with autism and their MA-matched controls.

Studies that have examined the drawing process, rather than just the end product, have provided evidence for unusual drawing styles in ASD, possibly indicative of weak central coherence. Fein, et al. (1990), for example, found that children with autism (5 to 17 years) would more often show fragmentation and overlap of parts in their drawings of a person than developmental-level matched control children. A failure to integrate the parts of the figure into a coherent whole was only ever shown by the children with autism and was taken as a sign of over-focused attention to detail.

Mottron and colleagues analysed the drawing style of a savant artist with Asperger syndrome (Mottron & Belleville, 1993) and a group of high-functioning adolescents and adults with autism (Mottron, Belleville et al., 1999). They observed in both studies that individuals with autism tend to draw the internal details first rather than the outline shape. This was considered as evidence for a local bias in processing visual information in autism. Booth et al. (2003) also found that boys with ASD more often showed a detailed-focused drawing style than typically developing boys and those with ADHD. The ASD boys were more likely to begin their drawing with a local detail, draw in a piecemeal fashion, and create less coherent drawings. This unusual feature-based drawing style was found to be specific to boys with ASD, and not related to poor executive control such as the ability to plan ahead in their drawing.

Several studies have employed the standard Rey figure as a means to study the drawing abilities and styles of individuals with ASD (see Table 7.1). Although the direct copy of the Rey figure has found to be comparable in individuals with ASD and matched controls, the recall of the figure is often deficient in ASD (Minshew & Goldstein, 2001; Prior & Hoffmann, 1990). This finding has been attributed to poor organisation in the original copy, such as taking a fragmented approach and starting with local details, rather than the global structure. Jolliffe and Baron-Cohen (1997) considered their modified Rey figure might have been oversimplified to the extent that ceiling effects pertained and a local bias could not be demonstrated by group differences.

Table 7.1 Published studies using the Rey figure in ASD

Reference	Participants (N, age range, mean FIQ/VMA)	Main findings
Steel, Gorman, & Flexman, 1984	1 x autism, 29 years, FIQ 91	Case study of mathematical savant with autism. Primarily focused on drawing internal elements of Rey figure; showed poor recall.
Prior & Hoffmann, 1990	12 x autism, CA 10-17, MA 8-17, FIQ 88 12 x TD, CA-matched, FIQ 100 12 x TD, MA-matched, FIQ 107	ASD < MA- & CA-matched controls on recall of the Rey figure, although equivalent in copy. Drawing process observed to differ between ASD and controls: ASD tended to draw details before the overall shape; recall often consisted of discrete components rather than the outline.
Rumsey & Hamburger, 1990	10 x autism males, CA 18-39, FIQ 96 15 x dyslexia males, CA 18-28, FIQ 104 25 x TD, CA 18-39, FIQ 107	No group differences in copy of the Rey figure, recall condition not administered.
Jolliffe & Baron-Cohen, 1997	17 x HFA, CA 19-46, FIQ 105 17 x Asperger syndrome, CA 18-49, FIQ 106 17 x TD, CA 18-49, FIQ 106	Used a modified Rey figure that consisted only of straight lines in order to clearly define what constitutes a global outline vs. a local internal feature. Participants viewed figure for 60 sec then asked to draw it from memory. Trend for ASD to draw in a fragmented way and use more lines cf. TD. 27% of ASD participants began with an internal feature, cf. 7% of TD.
Minshew & Goldstein, 2001	52 x HFA, CA 12-40, FIQ 93 40 x TD, CA 12-40, FIQ 97	HFA < TD in number of elements drawn in immediate and delayed recall trials. Groups were comparable in copy condition. Following 30-minute delay HFA retained 51% of elements, cf. 64% for TD.
Ropar & Mitchell, 2001	19 x autism, CA 9-18, VMA 6; 11 11 x Asperger syndrome, CA 8-15, VMA 9; 11 20 x MLD, CA 9-14, VMA 6; 11 19 x TD, CA 7-8, VMA 8; 0 18 x TD, CA 10-11, VMA 11; 6	ASD > MLD in Rey copy condition, no difference in recall. 55 % Asperger syndrome & 33% autism used a global drawing strategy (not differ from TD & MLD controls) No group differences for recognition of local details and global shape of the Rey figure. Some degree of meaning-imposition was used by TD, MLD, & autism (e.g., turning a feature into a nameable object). This was not observed in Asperger syndrome.

Task Development

To examine the effect of meaning on the recall of a geometric figure, performance on the Rey figure was contrasted with that on a novel Pram figure (designed by the author, see

Figure 7.2). The Pram figure was a line drawing of a stylised, but recognisable, everyday object of similar complexity to the Rey figure. The figure was structured around a main trapezium shape, with a vertical and horizontal line intersected by two diagonal lines, and included a variety of internal features. The hierarchical structure of the Pram figure was therefore considered comparable to the Rey figure. Furthermore, effort was made to equate the number of elements, both local and global, between the two figures (see Appendix K).

The Rey and Pram figures were used in the present study to measure drawing style and visual organisation, and to examine the effect of meaning on drawing accuracy. Participants were asked to copy each figure on separate occasions and, without prior warning, reproduce it from memory following a delay of 20 minutes. Drawings were scored for accuracy as well as organisational approach. Two independent indices of weak central coherence in drawing style were assessed: order of construction of local and global elements, and drawing style (continuous versus fragmented). These indices were similar to those devised by Booth et al. (2003) and adapted by the author for use with the Rey and Pram figures. The Pram figure was piloted on 32 TD children (7 to 15 years) to ensure feasibility and approximate equivalence in difficulty to the Rey figure. It was confirmed that both global and local strategies to copy and recall the Pram figure were evidenced in the drawing styles of the pilot participants.

Predictions

It was predicted that the extent to which recall accuracy benefits from the meaningfulness of the design will vary between individuals, and reflect individual differences in processing style. The greater advantage in recalling the Pram figure, relative to the Rey figure, is therefore indicative of strong central coherence. In terms of drawing style, it was predicted that weak central coherence would be shown by a preference for a local strategy in the construction of both the Rey and Pram figures. Strong central coherence, in contrast, would predict greater use of continuous drawing strokes, and relatively more global than local elements drawn in the initial stage of the drawing. In line with the findings from the TD literature, use of a global strategy is predicted to increase with age and intellectual ability in the TD sample.

7.2.2 Method

Materials

The Rey figure was reproduced so that the base rectangle measured 8 x 5.5cm, while the Pram figure measured 11cm across the top of the trapezium and the main vertical line measured 8.5cm. Both figures were printed in black on white A4 card (21 x 29.7cm). On each copy and recall trial, participants were provided with a piece of white A4 paper and a soft lead pencil. The Rey and Pram figures are shown in Figure 7.1 and Figure 7.2 respectively (approximately 50 percent of actual size).

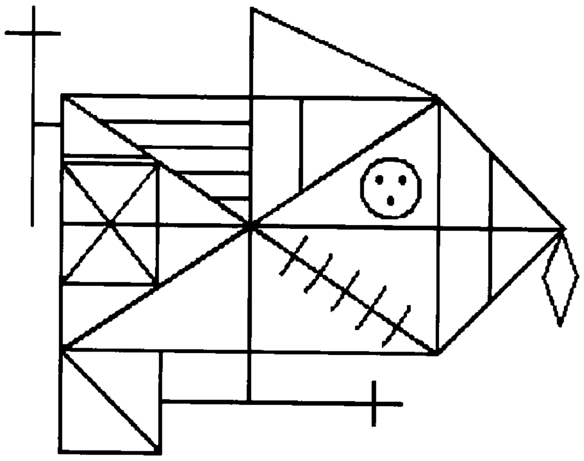


Figure 7.1 The Rey Complex Figure (Osterrieth, 1944; cited in Lezak, 1995).

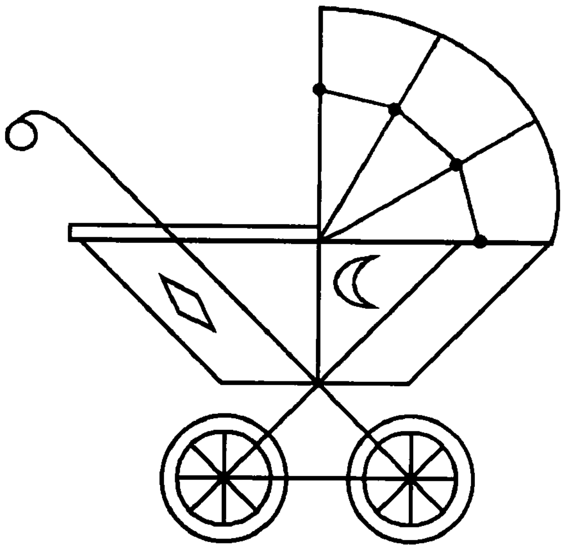


Figure 7.2 The Pram Figure.

Procedure

(1) Copy trial

The Pram and Rey figures were administered in two different test sessions, with the order of administration counterbalanced across participants. In the copy trial, the stimulus card of the figure was first presented. Participants were provided with a blank sheet of paper and a pencil and were asked to copy the figure, with the instruction to, "*try to make your drawing as much like mine as possible.*"

The researcher made a detailed record of the participant's copying sequence by drawing exactly what the participant drew on a separate page. Using between 10 and 15 colour pencils, the researcher sequentially changed colour when an element was completed or when the participant moved on to draw another part of the figure. Arrows were also drawn to indicate the direction of the drawing process. The order of the colours used was noted on the side of the page by the researcher. This method was a variant of the standard procedure detailed by Lezak (1995) whereby the participant is asked to change colour pencil during their drawing. It was considered that this approach could be disruptive for participants and affect their natural drawing style. The drawing process was videotaped for later analysis (if needed).

The researcher monitored the participant to ensure they did not rotate the stimulus card or attempt to trace the figure. If the participant said they had made an error in their drawing, the researcher encouraged the participant to cross it out and then continue. No time limit was imposed on the copy and the researcher made no mention of the recall requirement of the task.

(2) Delayed recall

The delayed recall trial was administered 20 minutes after the participant had finished his/her copy of the figure. Three or four tasks were administered during this delay period; the intervening tasks generally required verbal skills in order to avoid interference (see Appendix C for task administration order). Participants were then presented with a blank sheet of A4 paper and the lead pencil and were instructed to draw the picture they had copied a short while ago. Participants were told to try their best to draw as much as they could remember of the picture. As in the copy trial, the researcher copied the participant's drawing process by using coloured pencils to record the order and manner in which the figure was constructed. No time limit was imposed on participant's drawings.

7.2.3 Scoring: Rey Figure

Accuracy

The standard scoring system developed by L. Taylor (Spreen & Strauss, 1998) was used to assess the drawing accuracy of the Rey figure. This scoring system has been used with children (6 to 16 years) as well as adults and was therefore considered applicable for the present study. The scoring system divides the Rey figure into 18 separate elements (see Appendix K, Table K.1), which are scored individually for accuracy and placement. If an element is drawn accurately and placed correctly a score of 2 is awarded. One point is given if the element was drawn accurately, but placed incorrectly; or conversely, drawn incorrectly but placed correctly. If the element is recognisable, but not drawn or placed correctly, a score of 0.5 is assigned. If the element is omitted no points are awarded. The possible range of accuracy for the copy and recall trials of the Rey figure was 0 to 36.

Central coherence indices

(1) Order of construction

The order of elements drawn on copy and recall trials for each participant was taken from the on-line colour coding recorded by the researcher. Partially drawn elements were included in the construction order as it was considered this was an indication that the participant had attended to the element within the figure, even though it was incomplete.

An objective measure of the order in which the elements of the Rey figure were drawn was calculated. Of interest was the relative number of global features as opposed to local features drawn in the initial stage of construction. To achieve this, the 18 elements of the Rey figure were first categorised as belonging to one of four types: (1) the global external structure, (2) the global internal structure, (3) a local perimeter element, and (4) a local internal element (see Appendix K, Table K.2). Different weightings were assigned to each category that reflected its importance to the fundamental organisation of the figure: a global external element was awarded four points, a global internal element three points, a local perimeter element one point, and a local internal element zero points. The first third of elements drawn were assigned a respective weighting score. The average of these weightings was taken as the *order of construction* index, which could range from 0 to 4 points.

(2) Style

Six components of the Rey figure were rated for drawing style. The components are shown in separate colours in Figure 7.3. Style was defined by the degree of continuity in the drawing process. Each component was rated on a three-point scale, with two points being given where lines were drawn in a continuous stroke or drawn consecutively, one point where the component was partially fragmented or drawn separately, and zero points where the line was clearly disjointed in appearance or drawn in a piecemeal manner.

The participant's drawing process, as documented by the researcher at the time of drawing, was used to score the style index. The rating was based on the first drawing of a line and was made independent of accuracy; that is, if the component was partially drawn but recognisable, the rating was based on the lines that were present. If the component was absent or not recognisable, no rating was given. The final *style* index was therefore an average of all ratings for components present, and could range from 0 to 2.

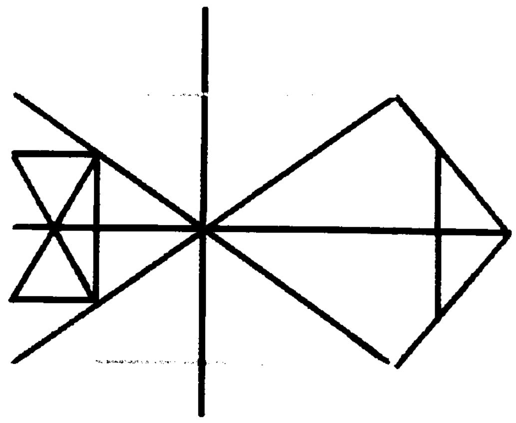


Figure 7.3 The six components scored for drawing style on the Rey figure.

7.2.4 Scoring: Pram Figure

Accuracy

A scoring system was devised by the author to obtain a measure of accuracy in drawing the Pram figure that was analogous to that of the Rey figure. The figure was initially divided into 16 discrete elements (see Appendix K, Table K.3). Each element was scored on a three-point scale according to how accurately it was drawn and placed: two points were given if both aspects were correct, one point if only the element or placement was correct, 0.5 points if the element was inaccurately drawn but recognisable and not in the correct position, and zero points if the element was not present. The range of accuracy for each drawing of the Pram figure was 0 to 32. The full detailed description of the scoring criteria for each element can be found in Appendix K (Table K.4).

Central coherence indices

(1) Order of construction

The order of construction of the Pram figure was quantified using the same procedure as the Rey figure. Initially the 16 elements of the Pram figure were classified into one of four hierarchical levels: (1) the global external structure, (2) the global internal structure, (3) a local perimeter element, and (4) a local internal element (see Appendix K, Table K.5). As it was common for participants to divide the vertical midline, the left diagonal line and the right diagonal line in their drawings, these three elements were reclassified into separate parts. A total of 20 elements were therefore coded for order of construction. As with the Rey figure, the first third of total elements drawn by the individual were assigned a score that represented in the degree of globality within the Pram figure (four points for global external elements, three points for global internal elements, one point for local perimeter elements, and zero points for local internal elements). The *order of construction* index was taken as the average of these weighted scores.

(2) Style

Six components of the Pram figure were selected for ratings of drawing style. The components are represented by different colours in Figure 7.4. Each component was rated on a three-point scale to reflect the degree of continuity in the drawing process, ranging from continuous (two points) to highly fragmented (zero points). As with the Rey figure, the style rating was based on what was drawn by the participant, with the average rating taken as the *style* index.

All inter-rater reliability checks for scoring accuracy, construction order, and style for the Rey and Pram figures are reported in Appendix K.

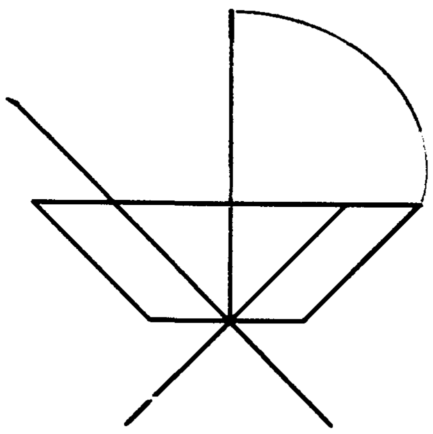


Figure 7.4 Six components scored for Style on the Pram figure.

7.2.5 Typical Development Results

All TD participants were administered the copy and recall trials of the Rey and Pram figure. Of the 204 participants, 101 were administered the Rey figure in the first test session, while the remaining received the Pram figure. This group was composed of an equal number of males (n = 49) and females (n = 52). Within each age group, the order of administration of the figures was balanced ($\chi^2 = 1.99$, $p = .57$) and distributed equally across male and female participants (all $\chi^2 < .77$, $p > .38$).

Order effects were somewhat expected on recall accuracy as participants were more aware of the recall requirement in the second administration of the Drawing Task (i.e., several participants queried whether they were expected to redraw the figure in the second test session, although the researcher did not divulge this information). No effect of order was found on accuracy however for the whole sample (all $z < 1.69$, $p > .08$), within male and female participants (all $z < 1.85$, $p > .06$), and within each age group (all $z < 1.75$, $p > .07$). In light of these findings the effect of administration order was not considered further.

Accuracy

In order to aid comparisons, accuracy scores for the copy and recall trials of the Rey and Pram figures were converted to percentages and mean age group scores are presented in Table 7.2. As the accuracy data were not normally distributed in each condition, non-parametric statistical tests were employed and all outliers remained in the analyses.

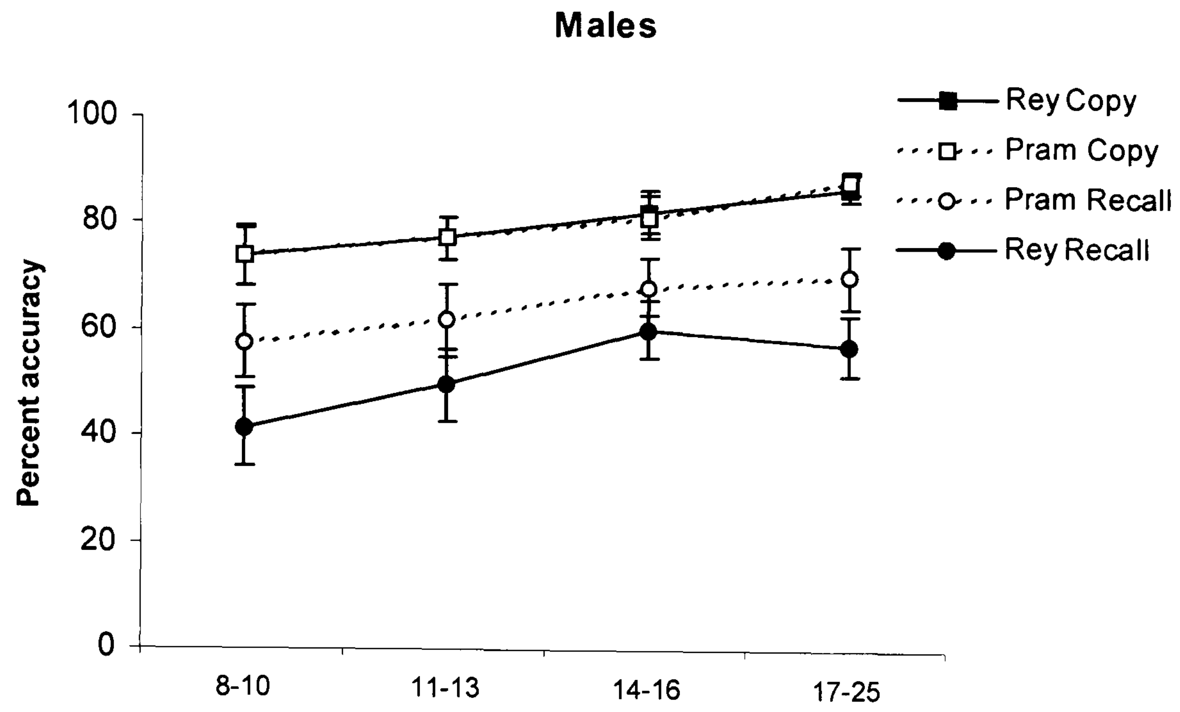
Table 7.2 Percent accuracy for the copy and recall trials of the Rey and Pram figures for each age group: Mean (*SD*).

Age group (years)	Rey figure		Pram figure	
	Copy[a]	*Recall*[b]	Copy[c]	*Recall*[d]
8-10	74.4 (12.1)	*43.9 (16.4)*	71.1 (10.8)	*56.8 (11.8)*
11-13	80.6 (9.9)	*52.6 (14.6)*	77.7 (9.8)	*64.2 (13.8)*
14-16	86.0 (10.3)	*61.0 (15.3)*	83.0 (10.1)	*69.9 (13.7)*
17-25	88.3 (6.5)	*57.4 (13.6)*	86.7 (6.9)	*70.3 (14.0)*

[a] 8-10 < 11-13, 14-16, 17-25, $p < .01$; 11-13 < 14-16, 17-25, $p < .004$
[b] 8-10 < 11-13, 14-16, 17-25, $p < .01$; 11-13 < 14-16, $p = .008$
[c] 8-10 < 11-13, 14-16, 17-25, $p < .004$; 11-13 < 14-16, 17-25, $p < .01$
[d] 8-10 < 11-13, 14-16, 17-25, $p < .008$; 11-13 < 14-16, 17-25, $p < .05$

Significant effects of age group were found in both copy and recall conditions of the Rey figure ($\chi^2 > 30.13$, $p < .0005$) and the Pram figure ($\chi^2 > 33.85$, $p < .0005$). Copy and recall accuracy significantly improved with age after 8-10 years, and after 11-13 years, while no difference in ability was observed between 14-16 years and 17-25 years. When analysing each gender separately, significant effects of age group were also found in the copy and recall conditions of the Rey figure (males $\chi^2 > 15.88$, $p < .002$; females $\chi^2 > 15.66$, $p < .002$) and the Pram figure (males $\chi^2 > 9.48$, $p < .03$; females $\chi^2 > 28.7$, $p < .0005$), although the pattern of age group effects differed between male and female participants. While a significant increase in copy and recall accuracy was found for females after 8-10 years (both Rey and Pram figures, all $z > 2.16$, $p < .04$), and again after 11-13 years for copy accuracy (all $z > 2.11$, $p < .04$), improvements in drawing accuracy for copy and recall were not observed until after 11-13 years for males (Rey between 11-13 and 14-16 years, $z > 1.95$, $p < .05$; Pram between 8-10 and 14-16 years, $z > 2.13$, $p < .04$, see Figure 7.5).

Male and female participants did not differ overall in their ability to draw the Rey and Pram figure, although a trend was apparent for male participants to copy the Pram figure more accurately than females ($z = 1.84$, $p = .07$).



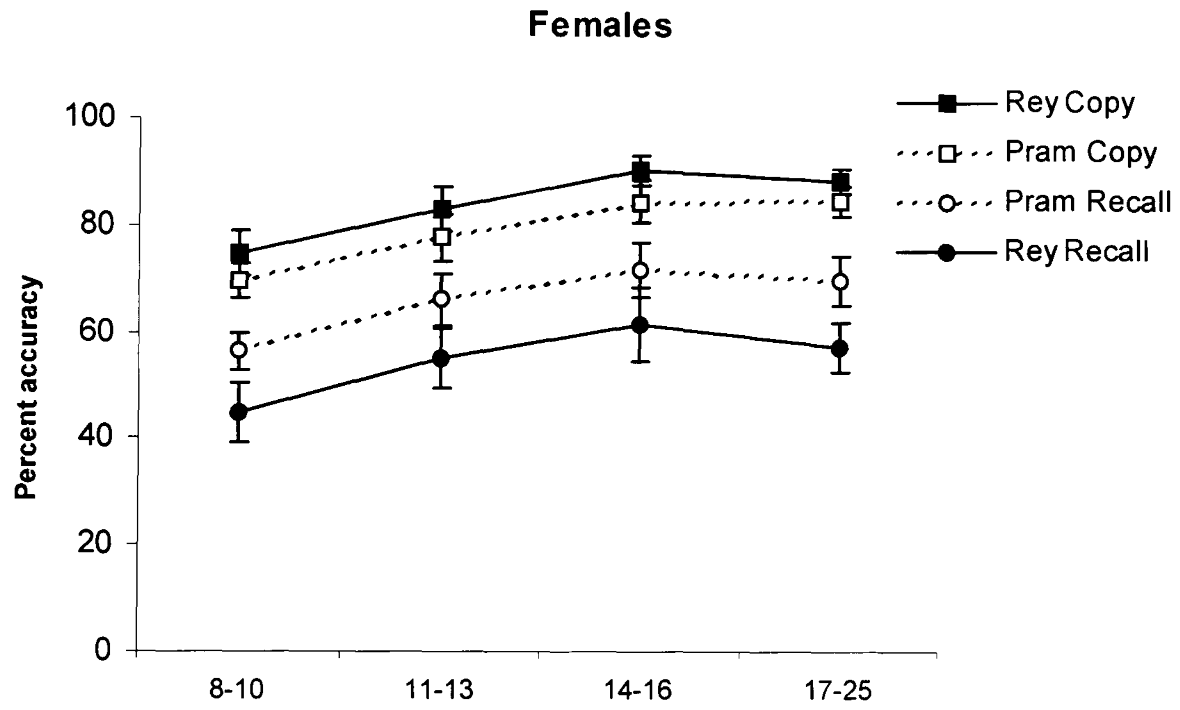


Figure 7.5 Percent accuracy for the copy and recall trials of the Rey and Pram figures for each age group for male and female participants. Error bars show 95% confidence intervals

The effect of meaning on copy accuracy

The Pram figure was designed to be of similar complexity to the Rey figure, differing only in meaningfulness. Despite the attempt to equate the two geometric forms, copy accuracy for the Pram figure was significantly lower than for the Rey figure over the whole TD sample (Wilcoxon's Signed Ranks Test, $z = 3.99$, $p < .0005$). This finding was significant in the two youngest age groups ($z > 1.96, p < .05$), but was not so evident in the two oldest age groups ($z < 1.64, p > .09$).

Comparisons between the Pram and Rey figure were also conducted by gender. The Pram figure was copied less accurately than the Rey figure for 73% of females (78 from

107; significant pairwise comparisons: $z = 5.05$, $p < .0005$). In contrast, this pattern of performance was found for 52% of males (50 from 97; $z = .20$, $p = .84$). The same result was found in each age group: male participants drew the Pram and Rey figure with equivalent accuracy (all $z < 0.75$, $p = .45$), while females consistently drew the Rey figure with greater accuracy than the Pram figure (all $z > 2.09$, $p < .04$). This gender effect in accuracy may have come about by females working more from an internal representation of a pram, rather than the existing model. This in turn might reflect stronger central coherence in females, or may be a function of prams having a strong gender stereotype.

The effect of meaning on recall accuracy

While the Rey figure was copied with greater accuracy than the Pram figure overall, the recall of the Rey figure was comparably lower than the Pram figure ($z = 10.42$, $p < .0005$). This effect was found in all age groups (all $z > 4.46$, $p < .0005$) and suggests that memory was weaker for non-meaningful than meaningful visual information. Participants of both genders also found it more difficult to recall the abstract than the meaningful design, with 85% of females and 81% of males drawing the Rey figure less accurately than the Pram figure at recall.

In order to obtain a measure of individual effects of meaning on recall accuracy, difference scores were calculated between the recall accuracy of the two figures (Pram minus Rey) for each individual. The difference score was normally distributed and allowed parametric statistical analyses to be performed. A two-way ANOVA did not find a significant main effect of age group ($F_{(3,196)} = 1.45$, $p = .23$), gender ($F_{(1,196)} = 0.35$, $p = .55$), or interaction of age group and gender ($F_{(3,196)} = 0.81$, $p = .49$).

Correlations: Benefit from meaning

The difference score showed no association with age in the TD sample ($r = -.03$, $p = .69$; males $r = -.13$, $p = .21$; females $r = .05$, $p = .63$). No reliable correlation was observed between any IQ measure and the difference score in the whole TD sample ($r = -.05$ to $.07$, $p > .29$) or TD females ($r = -.15$ to $-.10$, $p > .12$). A positive correlation was found between VIQ and the difference score in male participants ($r = .21$, $p = .04$), but not on any other IQ measure ($r = .002$ to $.17$, $p > .08$). Fisher's z transformation found that the magnitude of the coefficients differed significantly between genders (higher for males) for FIQ, VIQ and FIQ-BD (all $z_{r1-r2} > 1.91$, $p < .05$).

Coherence index

Although order of construction and style were logically independent (i.e., starting with local detail does not necessarily result in a fragmented drawing), the strong association between the two measures ($r_s = .15$ to $.45$, $p < .04$; $pr_{\text{age FIQ-BD}} = .13$ to $.38$, $p < .07$) suggested that they tapped the same underlying construct. A composite score was therefore constructed by summing the proportion scores for style (originally out of 2) and construction order (originally out of 4) for the copy and recall conditions of the Drawing Task. These scores were averaged across the Rey and Pram figures for conciseness (see Table 7.3). Note that higher scores indicate a more coherent drawing style. The copy coherence index for one 17-year-old male (FIQ 119, VIQ 114, PIQ 120) was deemed an outlier for his age group (copy = 0.62, recall = 1.06). His data were removed from analysis and assumptions of normality were subsequently met in each age group.

Coherence index data were analysed with a two-way repeated measures ANOVA with age group and gender as between-subjects factors, and condition (copy vs. recall) as the within-subject factor. There was a main effect of age group ($F_{(3,195)} = 15.42$, $p < .0005$), with oldest participants showing a more coherent drawing style compared to the younger age groups (Tukey's HSD, all $p < .003$). A significant main effect of gender was found ($F_{(1,195)} = 9.44$, $p = .002$), with females showing a more coherent drawing style than males. No interaction between gender and age group was detected ($F_{(3,195)} = 0.17$, $p = .92$; see Figure 7.6).

A main effect of condition was found ($F_{(1,195)} = 77.26$, $p < .0005$), with participants showing a more coherent drawing style in the recall compared to the copy condition. There was no interaction between condition and age group ($F_{(3,195)} = 1.43$, $p = .24$), condition and gender ($F_{(1,195)} = 1.34$, $p = .25$), or between condition, age group and gender ($F_{(3,195)} = 1.56$, $p = .20$). The main effect of condition was therefore consistent across age groups, and across male and female participants within each age group.

Table 7.3 Coherence indices (maximum = 2) for the copy and recall trials of the Drawing Task (averaged across Rey and Pram figures) for each age group: Mean (*SD*).

Age group (years)	Copy[a]	Recall[b]
8-10	1.21 (0.17)	1.30 (0.15)
11-13	1.22 (0.13)	1.28 (0.15)
14-16	1.26 (0.15)	1.36 (0.13)
17-25	1.37 (0.17)	1.44 (0.15)

[a]8-10, 11-13, 14-16 < 17-25, $p < .002$
[b]8-10, 11-13, 14-16 < 17-25, $p < .04$

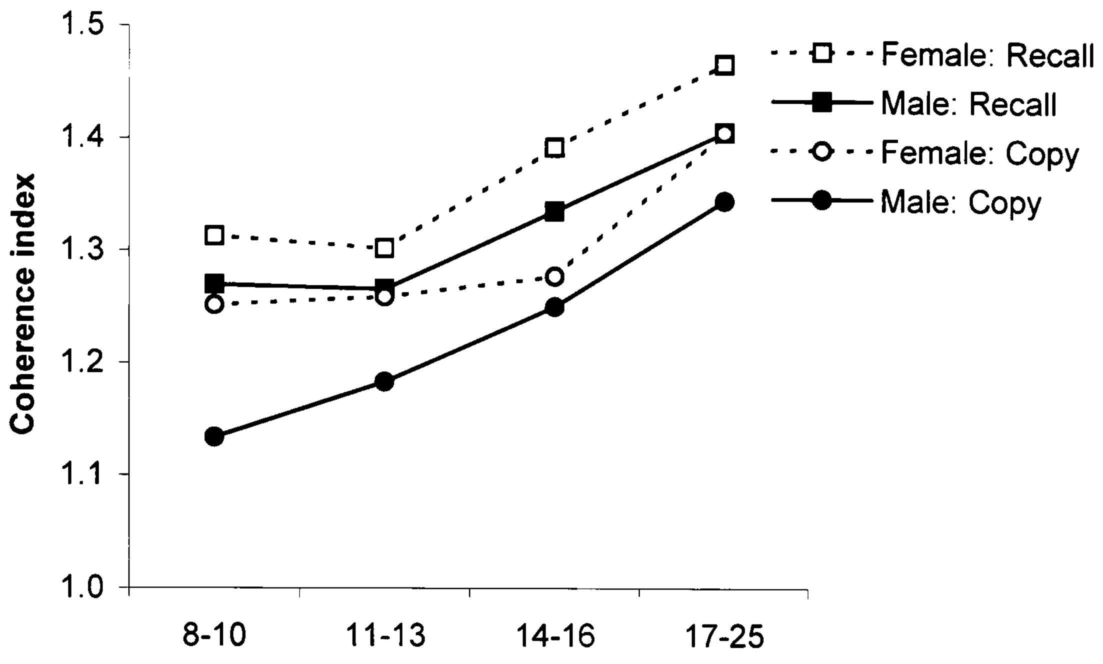


Figure 7.6 Mean coherence index (max = 2) for the copy and recall conditions of the Drawing Task for male and female participants in each age group.

Correlations: Coherence index

Examination of the data suggested that participants tended to distort the Pram figure in recall. It was considered that participants might have drawn on their internal representation of a canonical pram form, which did not match the present figure nor lend itself to the scoring criteria designed for the original Pram figure. For this reason, the coherence index for the copy condition (averaged across Rey and Pram figures) was selected as the main variable to study individual differences in drawing style.

The copy coherence index showed a strong positive correlation with age in the TD sample ($r = .40$, males $r = .49$, females $r = .36$, all $p < .0005$). While a significant positive correlation was found between the coherence index and FIQ ($r = .15$, $p = .04$) and PIQ ($r = .19$, $p = .006$), this association was not found with VIQ ($r = .08$, $p = .27$), or FIQ-BD ($r = .07$, $p = .30$). When analysing by gender, the correlation between FIQ and the coherence index fell below significance for both male ($r = .19$, $p = .06$) and female participants ($r = .15$, $p = .12$). The correlation between PIQ and the coherence index remained significant for males ($r = .25$, $p = .01$), but not for females ($r = .18$, $p = .07$). Furthermore, no reliable correlation was found between VIQ or FIQ-BD and the coherence index in either gender (males $r = .11$ to $.12$, $p > .25$; females $r = .07$ to $.08$, $p > .39$).

7.2.6 *ASD and Control Results*

The Rey and Pram figures were administered to all participants in the ASD and control groups. One 16-year-old ASD participant (FIQ = 94, VIQ = 107, PIQ = 81) was not able to reproduce any elements of the Rey figure from memory despite copying the figure with high accuracy (86%). The copy of the Rey figure for this participant was removed from analysis, while the copy and recall of the Pram figure remained[7].

Approximately half of the participants in each group received the Rey figure in the first test session (16 ASD, 15 control), while the remainder were administered the Pram figure. No order effects were found on the measure of drawing accuracy within each group (all $z < 1.13, p > .25$).

Accuracy

The percent accuracy scores for the copy and recall trials of the Rey and Pram figures are presented in Table 7.4. As the accuracy data for each group were not normally distributed, non-parametrical statistical analyses were used throughout. A significant group difference was found in recalling the Rey figure, with the ASD group performing significantly lower than the control group ($z = 1.99$, $p = .05$, one-tailed, see Figure 7.7). This difference was evident within high IQ participants ($z = 2.23$, $p = .03$), but not when comparing low IQ participants ($z = 0.32$, $p = .75$). No further group differences were detected overall, or between low and high IQ groups (all $z < 1.52, p > .12$).

Table 7.4 Percent accuracy for the copy and recall trials of the Rey and Pram figures for each group: Mean (*SD*).

		Rey figure		Pram figure	
	Group	Copy	*Recall*[a]	Copy	*Recall*
All	ASD	72.9 (19.5)	*39.0 (17.8)*	72.2 (15.5)	*54.3 (19.8)*
	Control	76.3 (15.1)	*48.4 (19.4)*	75.6 (16.3)	*59.0 (19.2)*
High IQ	ASD	77.5 (15.8)	*41.8 (18.2)*	75.8 (11.7)	*60.4 (16.1)*
	Control	81.2 (8.7)	*53.2 (17.2)*	80.6 (10.1)	*64.7 (15.4)*
Low IQ	ASD	52.3 (22.2)	*28.0 (11.0)*	57.3 (21.2)	*29.2 (12.9)*
	Control	55.8 (19.5)	*28.2 (15.5)*	54.7 (21.4)	*35.4 (15.5)*

[a]ASD < Control, $p = .05$; $ASD_{high} < Control_{high}$, $p = .03$

7 This participant copied the Pram figure (administered in the first session) with 78% accuracy and recalled it with 47% accuracy. As the removal of the Pram figure data for this participant did not significantly alter group means, his data remained in the analysis. The effect of meaning on drawing style could not be examined for this participant however.

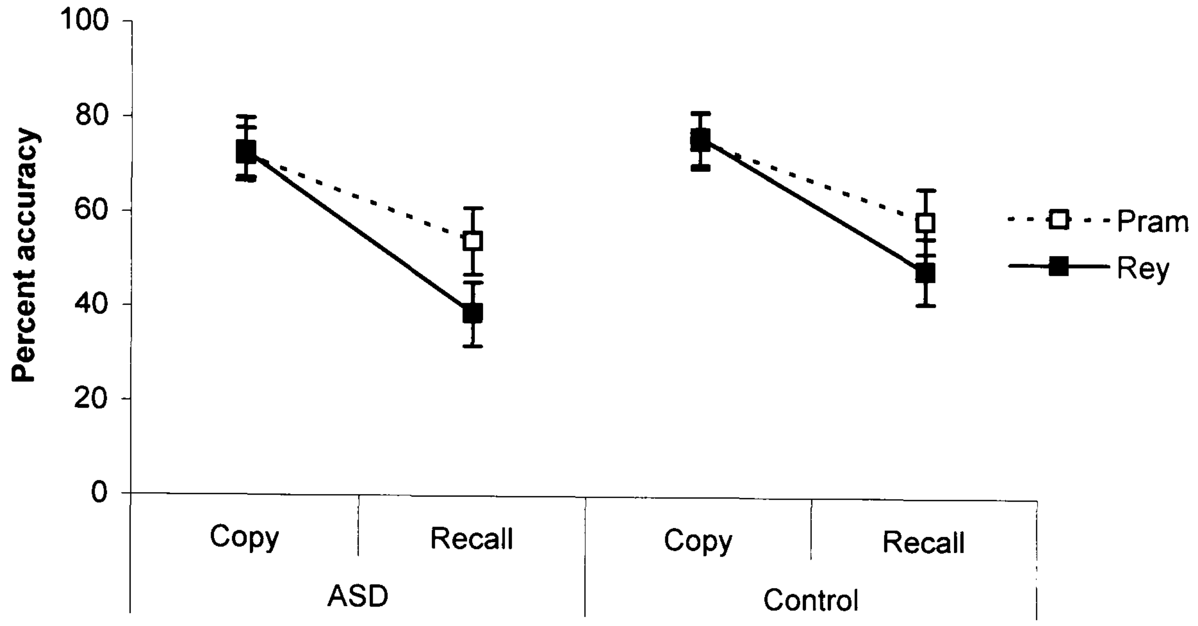


Figure 7.7 Percent accuracy for the copy and recall trials of the Pram figure and the Rey figure for each group. Error bars show 95% confidence intervals.

The effect of meaning on copy accuracy

Performance was contrasted between the Pram figure and the Rey figure to investigate the effect of meaning on drawing accuracy. The relative difficulty of copying the Rey figure was equivalent to copying the Pram figure for participants in both groups (ASD $z = 0.48$, $p = .63$; control $z = 0.33$, $p = .74$). Groups of high IQ and low IQ participants also did not differ in drawing accuracy between the two figure types (all $z < 1.29$, $p > .19$), which mirrors the finding in the TD sample where male participants drew the Pram and Rey figure with equivalent accuracy.

The effect of meaning on recall accuracy

The majority of participants in the ASD group (83%) and the control group (77%) were more accurate in recalling the Pram figure than the Rey figure ($z = 3.94$ and $z = 4.04$, respectively, $p < .0005$). This was also found within groups of high IQ (ASD $z = 4.09$, control $z = 3.79$, $p < .0005$), but not within groups of low IQ (ASD $z = 0.11$, $p = .92$; control $z = 0.94$, $p = .35$). The typical benefit from meaning in the recall of a geometric form was therefore not evident in groups of low-functioning individuals, from both ASD and control groups. Recall accuracy was low in both groups however (28% to 35% of elements correctly drawn), and it is possible that floor effects obscured any difference between meaningful and non-meaningful figure types.

As in the TD sample, difference scores were calculated between the recall accuracy of the two figures (Pram minus Rey) to obtain a measure of an individual's benefit from the meaningfulness in recall of the design. Opposite to predictions of weak central coherence theory, the mean difference score was higher in the ASD group ($M = 15.6$, $SD = 17.1$) than the control group ($M = 10.6$, $SD = 11.0$), although the groups did not differ statistically

($t_{(49.3)} = 1.33$, $p = .09$, one-tailed, unequal variances). This difference reached significance within groups of high IQ participants (ASD $M = 19.2$, $SD = 16.2$; control $M = 11.5$, $SD = 10.9$; $t_{(40.22)} = 1.94$, $p = .03$). The mean difference score was much reduced in the low IQ participants with ASD ($M = 1.2$, $SD = 13.4$), although not significantly different from that for low IQ controls ($M = 7.2$, $SD = 11.6$; $z = 0.96$, $p = .34$).

Correlations: Benefit from meaning

The difference in drawing accuracy between the Pram and Rey figures (indicating the effect of meaning) did not correlate with age (ASD $r_s = -.18$, $p = .35$; control $r_s = .02$, $p = .92$) or MA in either group (ASD $r_s = .20$, $p = .28$; control $r_s = .12$, $p = .53$). A positive correlation was found between the difference score and PIQ in the ASD group ($r_s = .37$, $p = .05$), with a trend for FIQ ($r_s = .34$, $p = .07$). Moderate correlations were also found between the difference score and VIQ ($r_s = .30$, $p = .11$) and FIQ-BD ($r_s = .26$, $p = .17$) in the ASD group, but neither reached significance. No association between the difference score and IQ was found in the control group ($r_s = .01$ to $.12$, $p > .51$), although the correlation coefficients were not found to differ in magnitude between groups (all $z_{r1\text{-}r2} < 1.40$, $p > .16$).

Coherence index

As conducted in the TD analyses, a coherence index was constructed for the copy and recall conditions (see Table 7.5) based on the summation of the proportion scores for the two coherence indices, style and order of construction, averaged across Rey and Pram figures. A higher score on the coherence index reflects a more coherent drawing style. As the coherence indices for the two groups were not significantly skewed (z-scores –0.76 to 0.36), parametric statistics were employed. No significant group differences were found between the ASD and control groups on the coherence index for copy ($t_{(60)} = 0.64$, $p = .26$, one-tailed), or recall conditions ($t_{(51.17)} = 0.13$, $p = .45$, unequal variances). Furthermore, no significant group differences were detected within groups of high IQ (copy $t_{(48)} = 0.58$, $p = .28$; recall $t_{(40.12)} = 0.06$, $p = .48$) or low IQ participants (copy $z = 0.48$, $p = .32$; recall $z < 0.01$, $p = .50$).

Table 7.5 Coherence indices (maximum = 2) for the copy and recall trials of the Drawing Task (averaged across Rey and Pram figures) for each group: Mean (*SD*).

	Group	Copy	Recall
All	ASD	1.21 (0.16)	1.33 (0.22)
	Control	1.24 (0.16)	1.33 (0.14)
High IQ	ASD	1.22 (0.17)	1.32 (0.20)
	Control	1.24 (0.14)	1.33 (0.13)
Low IQ	ASD	1.19 (0.12)	1.37 (0.27)
	Control	1.22 (0.23)	1.33 (0.19)

Correlations: Coherence index

The copy coherence index did not correlate significantly with age in either group (ASD $r = -.16$, $p = .38$; control $r = -.04$, $p = .85$). Similarly, the association between the coherence index and MA was not strong in either group (ASD $r = .01$, $p = .97$; control $r = .17$, $p = .37$). Positive correlations were observed between all IQ measures and the coherence index for both groups, but none reached significance (ASD: FIQ $r = .22$, $p = .25$; FIQ-BD $r = .14$, $p = .45$, PIQ $r = .34$, $p = .06$, VIQ $r = .08$, $p = .68$; control: FIQ $r = .29$, $p = .12$, FIQ-BD $r = .18$, $p = .33$, PIQ $r = .33$, $p = .07$, VIQ $r = .21$, $p = .26$).

7.2.7 *Summary and brief discussion of Drawing Task findings*

Drawing accuracy, in both copy and recall conditions of the Rey and Pram figures, increased with age in the TD sample (between 8-10 and 14-16 years). This improvement was found to occur earlier for females (between 8-10 and 11-13 years) than males (between 11-13 and 14-16 years). Recall accuracy was higher for the Pram figure than the Rey figure demonstrating that memory for meaningful information was better than non-meaningful information. This effect was found in the TD sample and in the ASD group, with no difference found between the ASD and control groups. Groups of low-functioning individuals (from both the ASD and control groups) did not show this benefit from meaning in recall however, which suggests a role for general intelligence in use of meaning to aid recall.

Age effects were found in the drawing style in TD, indicating that older participants took a more global drawing approach (i.e., drawing the outline or global features first and using continuous, non-fragmented lines) than younger participants. Female participants were also more global in their drawing style than male participants.

The ASD group was significantly less accurate in the recall of the Rey figure than the control group, which confirms previous findings in the autism literature (Minshew & Goldstein, 2001; Prior & Hoffmann, 1990). Poor recall of the Rey figure can be indicative of a piecemeal and disorganised approach to the copy of the figure, and it was predicted that differences in drawing style might be apparent between the two groups. The drawing process and product for the Pram and Rey figures was rigorously analysed between the ASD and control groups but no differences were found. This is counter to the predictions from weak central coherence theory and to previous studies that have proposed a local drawing style is characteristic of ASD (Booth et al., 2003; Jolliffe & Baron-Cohen, 1997; Mottron, Belleville et al., 1999). The present investigation took a very structured approach to scoring coherence in drawing style (e.g., unlike Booth et al.). Inspection of drawings from ASD participants did reveal examples of very fragmented drawing styles, and it is possible that the scoring system used in the present study was not sensitive enough to detect group differences. It remains to be determined whether a more subjective method of assessing drawing style (e.g., using a Likert scale for rater's impression of fragmentation) may have been more suitable.

The main index of processing style taken from the Drawing Task was the coherence index (incorporating both construction order and drawing style) for the copy condition of the Rey and Pram figures, with high scores indicating strong coherence. A second index of individual variability was the difference in recall accuracy between the two figure types (Pram minus Rey), which indicated the extent to which recall benefited from the meaningfulness of the design. High scores, again, were indicative of strong coherence. In order to capture both elements of coherence in drawing style (style and construction order), participants were classified as showing weak central coherence on two counts: (1) if they had more than one zero rating of style on the copy of the Rey and Pram figures (a zero style rating was only given when the element was clearly fragmented and drawn in a piecemeal manner); (2) if their mean order of construction score below 2.15 (max = 4.00; averaged across all conditions), indicating the tendency to draw local features in the initial stage of the drawing.

7.3 Picture Memory Task

The perception of a visual scene involves the appreciation of several individual objects and their organisation into a coherent structure. Much debate exists as to whether typically developing adults perceive the whole first and then decompose it into parts, or perceive the parts first, which become integrated into a whole (see Chapter 3). A widely held view is that visual scene perception is globally mediated and the identification of individual objects and surface features is secondary to perceiving the context (Aginsky & Tarr, 2000).

Empirical evidence has shown that semantic context plays an important role in the overall visual processing of pictures. When presented with briefly presented visual scenes, it has been shown that viewers can automatically extract information regarding the overall configuration or layout of the scene (Metzger & Antes, 1983). Furthermore, in object-detection studies it is commonly found that identification of an object is facilitated if consistent with the contextual scene (Biederman, Mezzanotte, & Rabinowitz, 1982; Hollingworth & Henderson, 1998; Palmer, 1975). Object identification is also facilitated in coherent as opposed to jumbled scenes even when the viewer is cued where look for the object (Biederman, 1972).

A processing style of weak central coherence would predict a disruption in the normal processes involved in perceiving a visual scene. Dawson et al. (2002) proposed that individuals with autism might show impairments in context memory as a consequence of weak central coherence. They suggest that a piecemeal processing style would result in details of a scene being stored as individual representations, rather than a coherent whole. Minor features may be accurately recalled, but alongside a lack of appreciation of the context.

In order to assess individual differences in processing a visual scene, a Picture Memory Task was developed for the present study. Participants were required to describe a scene after viewing it for 30 seconds. A recognition test followed to assess retention of surface form and details within the visual scene. Picture description is a commonly used procedure to elicit discourse from individuals (Lezak, 1995); responses can provide information about how effectively a participant perceives and integrates the elements of a visual scene. It is also advantageous as an open-ended task in which the natural processing style of the participant can be assessed.

Visual scene analysis in Typical Development

Research has shown that typically developing adults have an extremely large capacity and long durability of recognition memory for pictures (see Kobayashi, 1985, for a bibliography of picture memory studies). Despite this capacity, when analysing complex visual scenes, individuals tend to remember a meaningful interpretation of a picture, rather than the exact physical properties. Gernsbacher (1985), for example, found that memory for the orientation of a picture was found to decline rapidly in the recognition memory of undergraduate students, although memory for the general meaning of pictures remained high.

Metzger and Perlmutter (1984) investigated developmental differences in the local and global processing of pictures. Preschool children and young adults were compared in their ability to make same/different judgments of visual scenes that could differ in background and/or details. The background of the visual scene was found to be more salient than the details for both age groups as all participants found it more difficult to detect changes that appeared at the level of detail when no change occurred in the background. Processing style was found to differ by age, however. Older participants tended to process visual information in a sequential manner whereby global information was processed first, followed by the local information. Younger participants in contrast, would tend to process local and global information in a non-sequential manner before responding.

Developmental studies generally report an increase in the recall and recognition of visual information with age (e.g., Borges, Stepnowsky, & Holt, 1977; Mandler & Robinson, 1978; Pezdek, 1980), although some studies have found no such developmental trend (e.g., Bird & Bennett, 1974; Brown, 1973). It has been suggested that tasks that do not require mnemonic strategies for successful performance are not developmentally sensitive (Brown, 1973). The use of verbal strategies has been suggested to underlie the gender difference in favour of females on visual recognition tasks (McGivern et al., 1997; McGuinness & McLaughlin, 1982). Chipman and Kimura (1998), for example, found no gender difference in incidental memory for the content of complex scene, although a female advantage was found when easily labelled stimuli were used.

Visual scene analysis in ASD

As suggested above, weak central coherence theory predicts that individuals with ASD would have difficulty processing visual information in context and also in using context to facilitate recall. Ameli et al. (1988) found high-functioning individuals with autism could extract meaning from single objects to aid visual memory. Pring and Hermelin (1993) also demonstrated that artistically gifted individuals with autism could make semantic connections between items presented as pictures. It is plausible that weak coherence at the visual-semantic level may be more evident when analysing complex visual scenes consisting of several objects and thus requiring visual integration ability.

Recently Williams, Goldstein, and Minshew (2005) reported that high-functioning adults with autism were impaired on the immediate and delayed recall trials of the Family Pictures subtest from the Wechsler Memory Scale compared to matched controls. This measure assesses the ability to recall the actions and location of people in a scene as well as the relationships between people. Success on this test requires the ability to integrate elements in context and quickly identify the theme of the picture. Although the social content of the pictures may influence recall, the authors suggested the complexity of the visual stimuli contributed to the poor performance of the ASD group. Levy et al. (2005) further investigated this finding in a large number of high-functioning children with ASD (N = 121). Data from the Family Pictures subtest from the Children's Memory Scale were analysed but no impairments were found for memory of social scenes in their sample. Performance correlated significantly with real-life social and adaptive behaviour as measured by teacher report, which suggested that processing of the social content of the pictures might in part determine recall ability.

Jolliffe and Baron-Cohen (2001b) investigated visual processing in autism through a series of original tasks tapping the ability to integrate objects in context. In one task, high-functioning adults with ASD were compared with age- and IQ-matched controls in their descriptions of a drawing of an everyday scene. It was predicted that the individuals with ASD would tend to describe local or minor details rather than provide a coherent description of the scene. Although the ASD group said significantly fewer words than controls, the authors found no difference between the groups in the type of description given (e.g., initially describing characters, their actions, and the nature of the scene). Despite not showing a local descriptive style, the ASD group would occasionally give less coherent descriptions that lacked bridging statements. The ASD participants were also more likely to misinterpret events occurring in a scene, which was suggested to stem from a failure to integrate different objects within the picture.

Jolliffe and Baron-Cohen (2001b) also designed the pictures in their study to contain an object that was inappropriate for the context. Relative to the number of words said, participants in the ASD group were less likely to spontaneously comment on this incongruous object compared to controls. In a more directive task the ASD group made more errors and took longer to locate an out-of-place object when they were told it was present in a picture. A similar effect was found on another task where participants with ASD were less able to identify an incongruent object from a collection of objects that together would form a coherent scene (e.g., digging for treasure) relative to controls, even though they were able to identify the inconsistent object from a set of objects that formed a category (e.g., toys). The authors concluded that their results confirmed the existence of weak central coherence in autism at a conceptual level of visual processing.

Lopez and Leekam (2003, Experiment 1) investigated the use of visual contextual information in a priming paradigm. They found that, similar to age-matched controls, children with ASD were faster and more accurate in identifying single objects when they followed an appropriate contextual scene than after a neutral or inappropriate contextual scene. It was concluded that children with ASD did extract meaning from briefly presented (2 second) visual scenes, and these results were counter to Jolliffe and Baron-Cohen's (2001b) findings. Lopez and Leekam suggest that the cognitive demands differed between the two studies; determining the appropriateness of an object within a scene was more difficult than naming an object that followed a scene. Happé and Frith (2006) also suggest that priming from local elements of the scene may have occurred in the Lopez and Leekam study.

Task Development

In the Picture Memory Task participants were presented with a detailed picture of an underwater scene and were instructed to study the picture carefully. Following the removal of the picture, participants were asked to provide a description of what they had seen. This enabled analysis of recall style that could be classified as global (e.g., describing the overall setting, making a inference comment or personal response to the scene), as well as local (e.g., listing discrete objects, describing surface form or the position of objects).

A recognition test phase immediately followed the description. Pictures of single objects were presented one at a time to participants who were required to state whether the object was the same or different from that presented in the original scene. Test items consisted of: (1) identical objects from the underwater scene; (2) identical objects that were reversed in left/right orientation; and (3) objects related to an underwater scene but not present in the original drawing. Performance on the recognition test was taken as an indicator of surface form retention and local processing bias.

The design of the Picture Memory Task was initially piloted on 14 TD children, aged from 7 to 11 years. The initial exposure time of the underwater scene was changed from 20 seconds to 30 seconds, as it emerged that the short exposure made the recognition task too difficult for the younger participants (accuracy ranged from 40% to 55%, where chance = 50%). A total of 32 test items were piloted from which 25 were selected for the final task. Surface changes other than mirror reversals (e.g., colour, size) were piloted but proved to be very difficult to detect. As changes in left/right orientation produced a wide range of accuracy (25% to 75%), the final set of surface form changes included only items that had been reversed in orientation. This type of surface form change has been validated by Gernsbacher (1985) and can be considered analogous to changes in word order in sentence recognition studies, whereby the surface form is altered without affecting meaning. Test items consisting of objects related to the underwater scene were confirmed as adequate foils with accuracy rates ranging from 22% to 100% for individual items.

Predictions

It was predicted that weak central coherence would be manifest in descriptions of the underwater scene that focused on surface details and made relatively less mention of the overall theme or gist of the picture. Younger typically developing participants were predicted to show a more local style in their descriptions, while older participants were expected to show a more global recall style and begin their descriptions by identifying the type of scene, for example.

As memory studies have shown that surface information is typically lost more readily than semantic information (e.g., Gernsbacher, 1985), it was predicted that detecting changes in surface form (i.e., left/right orientation) would be harder than detecting changes in semantic form (i.e., different types of fish) in TD. Individuals with weak central coherence processing style are predicted to show greater proficiency on the recognition test overall through an enhanced tendency to attend to and retain detail information.

7.3.2 Method

Materials

A colour drawing of an underwater scene was used as the main stimulus in the Picture Memory Task (see Figure 7.8). The picture was created by the author from images extracted from the Clip Art library (Microsoft Office online) and comprised several objects that might be found in an ocean scene (e.g., deep-sea divers, corals, fishes, anchor, shipwreck, seahorse, jellyfish, shells, starfish, seaweed). The picture was designed to depict a typical underwater scene that did not require the appreciation of social cues. Stimuli from the test phase consisted of 25 single images of ocean-related objects. Fifteen images had been present in the main picture; five of these fifteen images were modified by reversing the orientation of the image from left to right. The final ten images were not present in the underwater scene but consisted of different forms of present items (e.g., a different type of fish) or associated objects to the underwater context (e.g., an octopus).

The presentation and timing of all stimuli was controlled by SuperLab Pro software run by a laptop computer. The underwater scene picture measured 32 cm by 24 cm (1024 x 768 pixels) and filled the 15-inch display of the touch screen. Individual test images appeared centrally on the touch screen in a set random order. The images were the same size as when presented within the underwater scene (which varied from 2 cm by 4 cm to 13 cm by 7 cm), while new items were comparable in size. All images for the test phase were presented in isolation on a neutral coloured background.

Figure 7.8 The underwater scene from the Picture Memory Task.

Procedure

Participants were presented with the underwater scene for 30 seconds and were asked to study it carefully. After the picture was removed from view, the participant was asked to describe what they had seen. Participants had 90 seconds to provide their description. The researcher gave one prompt (*"Is there anything else you can tell me about the picture"*) if a short description was offered. At the end of the 90 second period, the phrase "*Did the following appear in the picture*?" appeared on the screen and was also read aloud by the researcher. The 25 test images were presented individually and participants had unlimited time to make a "same" or "different" judgment to each image. On 40% of the trials a *same* test item was presented that had appeared in the original scene. The remaining 60% of trials consisted of *different form* (40%) and *mirror image* (20%) test items. It was emphasised to participants that each object had to look exactly the same as in the previous picture for an accurate "same" response.

Approximately 30 minutes after the completion of the recognition test, participants were asked to provide another description of the picture they had studied. One prompt was given by the researcher if necessary (e.g., "*The colourful picture that you saw a while ago*") and participants were given adequate time to give their answer.

Participants' descriptions of the underwater scene were audio recorded and later transcribed. The researcher recorded participants' verbal responses to same/different judgments made during the test phase.

Qualitative data analysis

Participants' descriptions of the underwater scene were examined at both global and local levels. The following four features were scored as "global" responses: (i) an overall or global description of the scene was given as the first response (e.g., *"it's an underwater picture," "it's a sea-bed"*); (ii) an overall description was ever used; (iii) an inference comment was made (e.g., *"they are looking for treasure"*); (iv) an emotional or personal response was made (e.g., *"I liked it"*) or a comment made referring to the quality of the picture (e.g., *"It looked like it was from a Disney film"*). The following three "local" features were measured: (i) the total number of different objects listed that had been present in the underwater scene (objects that were listed by the participant that were not present in the picture were also recorded and were classified as being related or unrelated to the scene); (ii) the total number of references made to surface detail (e.g., colour, shape, size, pattern, number) and the number of descriptive terms used (e.g., *"old sunken ship"*); (iii) the total number of references made to the position of objects on the screen (e.g., *"a crab in left corner"*) or in relation to one another (e.g., *"a lobster that was on a rock"*). Inter-rater reliability checks for local and global features are presented in Appendix L.

7.3.3 Typical Development Results

All 204 TD participants were administered the Picture Memory Task. Recognition test and immediate scene description data were available for all participants, while the delayed scene description was missing for two participants. This condition was not administered to one 10-year-old male, while one 15-year-old male was not able to provide any information about the picture following the 30-minute delay.

Scene Description: Global measures

Table 7.6 presents the percentage of participants in each age group who showed a global recall style as measured by the use of a global description, inference comment or personal response while describing the underwater scene. The chi-square test or Fisher's exact test were used to assess the difference in proportions across age group. Significant differences were found in the proportion of participants who began their description using an overall or global phrase, in immediate (χ^2 = 35.8, p < .0005) and delayed conditions (χ^2 = 23.0, p < .0005). Similarly, significant differences were detected in the proportion of participants in each age group who gave a global description at any point during their account of the underwater scene in the immediate ($\chi^2 = 25.5$, $p < .0005$) and delayed condition ($\chi^2 = 13.3, p = .004$).

The number of participants who used inference comments or personal responses was low across all age groups (see Table 7.6). A tendency was shown for inference comments to be of greater frequency in the youngest age group, however group comparisons did not reach statistical levels of significance (Fisher's exact test, all $p > .06$). Furthermore, the distribution of participants who gave at least one personal response to the picture was not significantly different across age groups (all $p > .30$).

Table 7.6 Mean percentage of participants in each age group who scored on each global measure in the immediate and delayed recall conditions of the Picture Memory Task.

Age group (years)	Global description at start (ever)				Inference comments		Personal responses	
	immediate[a]		*delayed*[b]		immediate	*delayed*	immediate	*delayed*
8 - 10	32%	(54%)	*36%*	*(53%)*	20%	*19%*	6%	*8%*
11 - 13	44%	(67%)	*44%*	*(65%)*	9%	*5%*	9%	*5%*
14 - 16	67%	(78%)	*60%*	*(74%)*	10%	*6%*	2%	*0%*
17 - 25	84%	(95%)	*79%*	*(84%)*	9%	*7%*	7%	5%

[a]start 8-10, 11-13 < 14-16 < 17-25; $p < .05$; ever 8-10 < 14-16 < 17-25; 11-13 < 17-25; $p < .01$
[b]start 8-10 < 14-16 < 17-25; 11-13 < 17-25; $p < .05$; ever 8-10 < 14-16, 17-25; 11-13 < 17-25; $p < .03$

A composite score was constructed to capture the degree of global recall style for each participant. One point was assigned for each of the four global measures: if global description made initially, if a global description was made ever, if an inference comment was made, and if a personal comment was made. The sums were averaged across immediate and delayed recall conditions and ranged from 0 to 4.

Developmental effects were found on the global composite ($\chi^2 = 22.56$, $p < .0005$), with a linear increase of global recall style with age group (8-10, 11-13 < 17-25, $p < .0005$; 8-10 < 14-16 < 17-25, $p < .03$). Similar effects of age group were found for male ($\chi^2 = 9.09$, $p = .03$) and female participants ($\chi^2 = 13.54$, $p = .004$; see Figure 7.9). No effect of gender was found overall ($z = 0.94$, $p = .35$), or within each age group (all $z < 1.23, p > .21$).

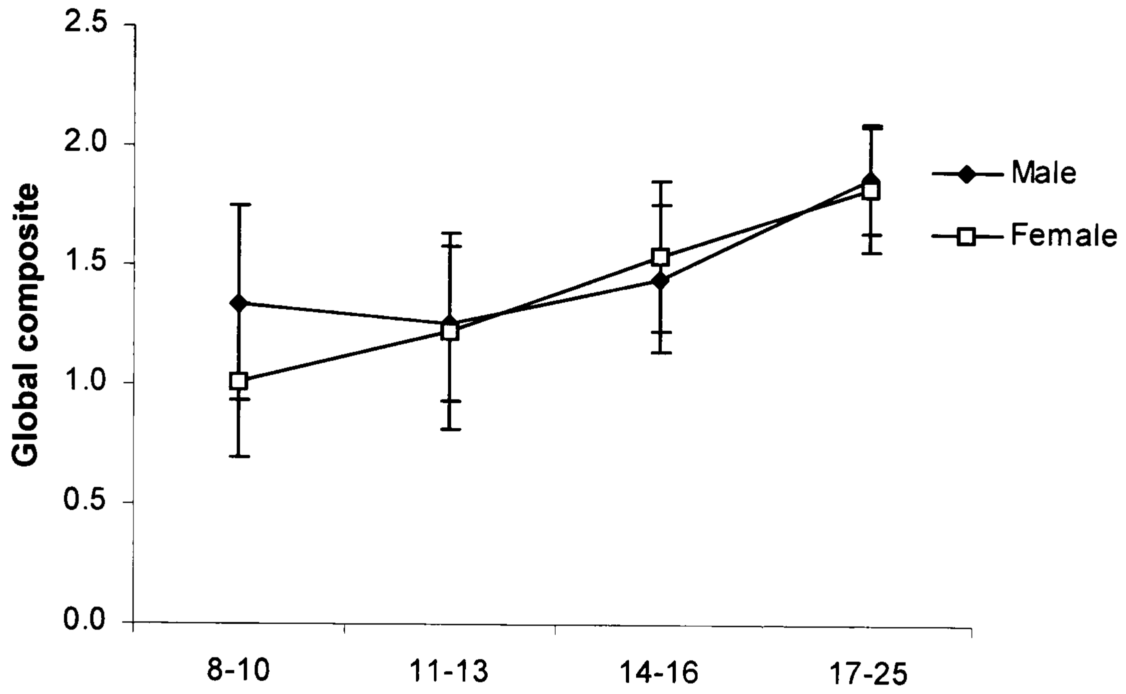


Figure 7.9 Mean global composite score for the scene description of the Picture Memory Task for male and female participants in each age group. Error bars show 95% confidence intervals.

Scene Description: Local measures

In order to capture a local recall style that was independent of recall ability, the number of references made to surface detail and position, relative to the number of objects listed, was calculated (see Table 7.7). As the data were extremely positively skewed for each age group (z-scores = 2.5 to 10.1) and contained outliers, non-parametric statistics were used. Significant age group effects were found for the relative number of references made to the position of objects, both in the initial ($\chi^2 = 9.49$, $p = .02$) and delayed ($\chi^2 = 12.59$, $p = .006$) description of the underwater scene. The relative number of references made to surface details was not significantly different between age groups in the immediate ($\chi^2 = 7.13$, $p = .07$) or delayed ($\chi^2 = 3.83$, $p = .28$) conditions.

Table 7.7 Number of references to surface detail and position of objects as a proportion of the number of objects listed by individuals within each age group: Mean (*SD*).

Age group (years)	Surface detail		Position	
	immediate	*delayed*	immediate[a]	*delayed*[b]
8 - 10	0.21 (0.20)	*0.28 (0.34)*	0.22 (0.35)	*0.25 (0.41)*
11 - 13	0.34 (0.52)	*0.28 (0.40)*	0.13 (0.19)	*0.15 (0.25)*
14 - 16	0.16 (0.24)	*0.18 (0.23)*	0.25 (0.32)	*0.33 (0.43)*
17 - 25	0.21 (0.21)	*0.24 (0.32)*	0.29 (0.31)	*0.36 (0.38)*

[a]8-10, 11-13 < 17-25; $p < .03$
[b]8-10, 11-13 < 17-25; 11-13 < 14-16; $p < .03$

Figure 7.10 presents the mean number of references made to surface detail (top graph) and object position (bottom graph), as a proportion of the number of objects listed (averaged across immediate and delayed conditions) for male and female participants in each age group. Female participants made significantly more references to surface details than males ($z = 2.02$, $p = .04$). This gender difference was not significant within any age group however (all $z < 1.35$, $p > .17$). In contrast, male participants made significantly more references to the position of objects than females ($z = 2.74$, $p = .006$), although this gender effect was only observed within the oldest age group ($z = 3.32$, $p = .001$).

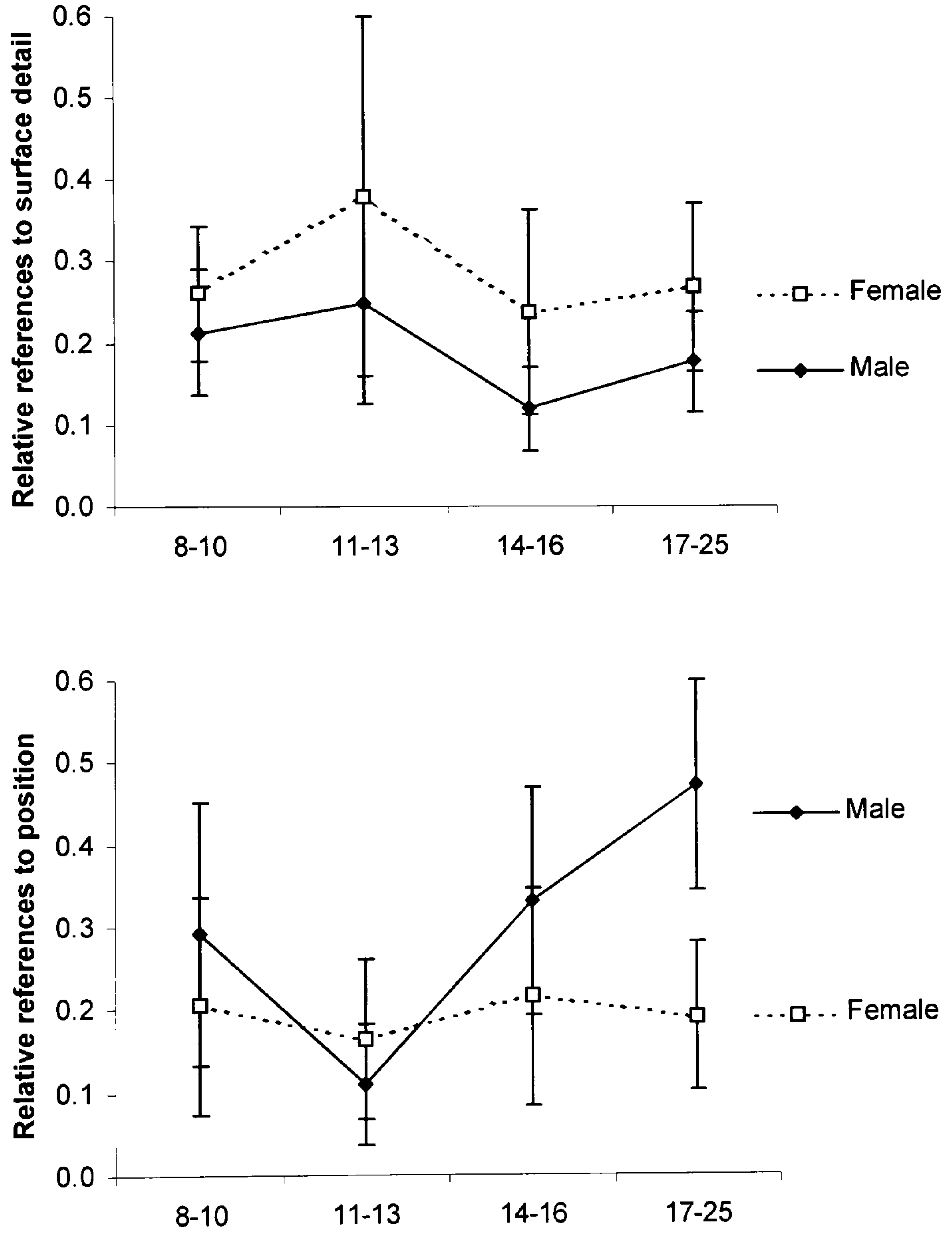


Figure 7.10 Number of references to surface detail and position of objects as a proportion of the number of objects listed (averaged across immediate and delayed conditions) by male and female participants within each age group. Error bars show 95% confidence intervals.

Significant effects of age group were found for male participants on the relative number of references to object position ($\chi^2 = 16.65$, $p = .001$). Males in the youngest age group made significantly fewer references to object position than the oldest age group

($z = 2.22$, $p = .03$). Males in the 11-13 year group also made significantly fewer references to object position than the two oldest age groups (all $z > 2.22$, $p < .04$). No significant effects of age group were found for male participants on the relative number of references to surface detail ($\chi^2 = 5.83$, $p = .12$). Furthermore, no significant age effects were found for female participants on either local style (all $\chi^2 < 1.12$, $p > .76$).

Correlations: Global measures

The global description composite showed a significant positive correlation with age ($r_s = .35$, males $r_s = .34$, females $r_s = .34$, all $p < .001$). A significant positive correlation was also found between the global description composite and IQ (FIQ $r_s = .22$, $p = .002$; VIQ $r_s = .19$, $p = .007$; PIQ $r_s = .20$, $p = .005$; FIQ-BD $r_s = .19$, $p = .006$). These correlations generally held for male participants (FIQ $r_s = .30$, $p = .003$; FIQ-BD $r_s = .32$, $p = .001$; VIQ $r_s = .33$, $p = .001$; PIQ $r_s = .18$, $p = .07$), but were only significant with PIQ in female participants (FIQ $r_s = .13$, $p = .17$; VIQ $r_s = .04$, $p = .65$; PIQ $r_s = .20$, $p = .04$; FIQ-BD $r_s = .06$, $p = .58$). The strength of the correlation coefficients differed between genders for VIQ ($z_{r1-r2} = 2.07$, $p = .04$).

Correlations: Local measures

The relative number of references to surface details, averaged across immediate and delayed recall conditions, showed no reliable correlation with age in the TD sample ($r_s = -.07$, $p = .30$; males $r_s = -.10$, $p = .34$; females $r_s = -.03$, $p = .76$). No reliable correlation between this marker of local recall style and intellectual ability was found (FIQ $r_s = -.04$, $p = .62$; FIQ-BD $r_s = .009$, $p = .90$; PIQ $r_s = -.05$, $p = .46$; VIQ $r_s = -.008$, $p = .91$; males all $r_s < .13$, $p > .23$; females all $r_s < -.15$, $p > .13$).

A second measure of local recall style was the relative number of references to the position of objects, averaged across immediate and delayed recall conditions. This measure correlated significantly with age in the whole sample ($r_s = .23$, $p = .001$), and in males ($r_s = .35$, $p < .0005$), but not females ($r_s = .14$, $p = .16$; $z_{r1-r2} = 1.57$, $p = .12$). A significant correlation was found between the second index of local recall style and IQ (FIQ $r_s = .30$; FIQ-BD $r_s = .27$; PIQ $r_s = .29$; VIQ $r_s = .25$; all $p < .0005$). The correlation between the local recall style and FIQ and PIQ was apparent in both genders (all $r_s > .23$, $p < .03$). Only in the male participants did VIQ and FIQ-BD correlate significantly with this marker of local recall style ($r_s > .29$, $p < .004$). These correlations did not reach significance in female participants ($r_s < .14$, $p > .17$), although the correlation coefficients did not differ significantly in strength between genders (all $z_{r1-r2} < 1.76$, $p > .08$).

Recognition test

The number (and percentage) of correct same/different judgments in the recognition test for each item condition is presented in Table 7.8 for each age group. For each age group, one-sample t-tests were performed to verify if mean detection performance differed from chance. All groups performed above chance when identifying *same* items (all $t > 7.55$, $p < .0005$) and *different form* items (all $t > 6.40$, $p < .0005$). The oldest age group did not differ from chance performance in determining *mirror image* items ($t_{(55)} = 1.45, p = .15$), while all younger age groups performed significantly *below* chance on the *mirror image* items (all $t > 2.96, p < .02$). This indicated that the majority of participants did not detect the change in orientation.

Table 7.8 Number and *percentage* of correct same/different judgments for the three test item conditions in the Picture Memory Recognition Test for each age group: Mean (*SD*).

Age group (years)	Same (max = 10)	Different	
		Form (max = 10)	Mirror Image (max = 5)
8 - 10	6.56 (1.38) *66%*	6.39 (1.60) *64%*	2.00 (1.05) *40%*
11 - 13	6.63 (1.41) *66%*	6.58 (1.37) *66%*	2.09 (1.06) *42%*
14 - 16	6.94 (1.35) *69%*	6.73 (1.39) *67%*	2.02 (1.16) *40%*
17 - 25	7.16 (1.30) *72%*	6.48 (1.58) *65%*	2.29 (1.11) *46%*

Signal detection analyses were conducted in order to take into account participants' decision bias on the recognition task. The mean non-parametric measure of sensitivity A' for each age group is presented in Table 7.9 for discriminations between *different form* and *same* items, and between *mirror image* and *same* items. As A' was generally negatively skewed for each age group (z-scores ranged from –0.85 to –5.45 for *different form* items, and –0.55 to –1.97 for *mirror image* items), non-parametric statistical tests were used.

Age groups were equivalent on the non-parametric measure of bias (B") for *different form* ($\chi^2 = 2.56, p = .47$) and *mirror image* detections ($\chi^2 = 0.86, p = .83$). Mean bias indices were negative for all age groups (ranging from –0.03 to –0.13 for *different form* items, and for –0.42 to –0.49 for *mirror image* items) indicating the tendency for participants to say the test item was present in the original scene. No gender effects were found on B" overall and within each age group (all $z < 1.43, p > .16$).

No significant effect of age group was found on A' for *different form* and *mirror image* test items ($\chi^2 = 5.17$, $p = .16$, $\chi^2 = 6.36$, $p = .10$, respectively). This was also found when including only male participants ($\chi^2 < 5.20$, $p > .15$) or female participants ($\chi^2 < 3.65$,

$p > .29$) in the analysis. No main effects of gender emerged in the whole group for recognition sensitivity for *different form* items ($z = 0.14$, $p = .89$) or *mirror image* items ($z = 0.41$, $p = .68$), nor within each age group (all $z < 0.80$, $p > .42$).

A' was significantly higher for *different form* items than *mirror image* items for the majority of participants (pairwise comparisons: all $z > 4.56$, $p < .0005$), whether male or female (all $z > 2.75$, $p < .007$) and characterised at least 75% of participants within each age group.

Table 7.9 Recognition sensitivity (A') for detecting different form and mirror image test items on the Picture Memory Recognition Test for each age group: Mean (*SD*).

Age group (years)	A' Different Form	A' Mirror Image
8 - 10	0.71 (0.13)	0.53 (0.17)
11 - 13	0.72 (0.15)	0.56 (0.20)
14 - 16	0.76 (0.11)	0.56 (0.22)
17 - 25	0.76 (0.11)	0.62 (0.20)

Correlations

A' for *mirror image* and *different form* items were averaged for each participant to provide a single index of discrimination ability for the correlation analyses. Mean A' significantly increased with age in the TD sample ($r_s = .20$, $p = .005$). When analysing by gender, this correlation was significant for male participants ($r_s = .23$, $p = .02$) and approached significance for female participants ($r_s = .17$, $p = .07$; $z_{r1\text{-}r2} = 0.44$, $p = .66$).

Recognition test performance was related to intellectual ability, with significant positive correlations found between mean A' and FIQ ($r_s = .23$, $p = .001$), FIQ-BD ($r_s = .15$, $p = .03$), PIQ ($r_s = .24$, $p = .001$), and VIQ ($r_s = .15$, $p = .03$). Correlations between mean A' and intellectual ability were generally stronger in males (FIQ $r_s = .23$, $p = .01$; FIQ-BD $r_s = .19$, $p = .06$, PIQ $r_s = .28$, $p = .006$; VIQ $r_s = .18$, $p = .08$) than females (FIQ $r_s = .18$, $p = .06$, FIQ-BD $r_s = .09$, $p = .34$, PIQ $r_s = .19$, $p = .05$, VIQ $r_s = .11$, $p = .25$), but did not differ reliably in strength between genders (all $z_{r1\text{-}r2} < 0.71$, $p > .48$).

7.3.4 *ASD and Control Results*

The Picture Memory Task was not administered to one 14-year-old boy from the ASD group who did not attend the final test session. A further 14-year-old boy from the ASD group responded "different" on 96% of the recognition test trials and his data were removed from analysis in this test phase as it was considered he was not performing the task. The groups remained well matched on age and all IQ measures after the exclusion of these two participants from analyses (all $t < 0.09$, $p > .93$).

Data for the immediate description of the underwater scene were available for 30 ASD participants and 31 control participants. In the delayed description of the scene, four participants with ASD and one control participant were not able to provide any information about the picture and the data were classified as missing. Again, the groups remained matched on age and IQ after the exclusion of these five participants (all $t < 0.57$, $p > .57$).

Scene Description: Global measures

The percentage of participants in each group who showed a global recall style as measured by the use of a global description (initial or ever), an inference comment or a personal response in the immediate and delayed description of the underwater scene is presented in Table 7.10. A similar proportion of participants in each group gave a global description as an initial response when first asked to describe the picture ($\chi^2 = 0.41$, $p = .52$) and following a delay of 30 minutes ($\chi^2 = 2.72$, $p = .10$). Furthermore, the proportion of participants who ever used a global description did not differ between groups for immediate ($\chi^2 = 0.84, p = .36$) and delayed conditions ($\chi^2 = 0.004, p = .95$).

The proportion of participants who made at least one inference comment regarding the underwater scene in their initial description did not differ significantly between groups (Fisher's exact test, $p = .47$). In the delayed description of the picture however, significantly more participants with ASD made an inference comment than control participants ($p = .04$). The proportion of participants making a personal response was not significantly different across groups in the immediate ($p = .35$) and delayed conditions ($p = .09$).

Table 7.10 Mean percentage of participants in each group who scored on each global measure in the immediate and delayed recall conditions of the Picture Memory Task.

	Group	Global description at start (ever)		Inference comments		Personal responses	
		immediate	*delayed*	immediate	*delayed*[a]	immediate	*delayed*
All	ASD	47% (63%)	*65% (69%)*	17%	*15%*	10%	*12%*
	Control	55% (74%)	*43% (70%)*	10%	*0%*	3%	*0%*
High IQ	ASD	54% (67%)	*73% (77%)*	21%	*18%*	13%	*9%*
	Control	56% (80%)	*46% (79%)*	4%	*0%*	4%	*0%*
Low IQ	ASD	17% (50%)	*25% (25%)*	0%	*0%*	0%	*25%*
	Control	50% (50%)	*33% (33%)*	33%	*0%*	0%	*0%*

[a] ASD > Control, $p = .04$; $ASD_{high} > Control_{high}$, $p = .04$

Group comparisons were also made including either high- or low-functioning participants. Low IQ ASD participants were less likely to use initially a global description when first asked to describe the underwater scene than low IQ control participants. Only one ASD participant from six used this response type compared to three from six in the control group, but Fisher's exact test did not show this group difference to be significant in a one-tailed test ($p = .27$). No further group differences were detected between groups of low IQ participants in their use of global descriptions, inference comments, or personal responses (all $p > .45$). There was a tendency for high IQ ASD participants to use a global description more often in the delayed description than control participants ($\chi^2 = 3.42$, $p = .06$). Moreover, the high IQ ASD participants used inference statements more than control participants in the delayed description condition (Fisher's exact test, $p = .04$). No further group differences were apparent between high IQ groups (all $p > .10$).

As in the TD sample, a global composite score was constructed consisting of all four global indices averaged across immediate and delayed recall conditions. No group differences were found on the global composite between all participants (ASD $M = 1.43$, $SD = 1.02$, control $M = 1.29$, $SD = 0.83$; $z = 0.68$, $p = .25$, one-tailed), those of high IQ (ASD $M = 1.60$, $SD = 1.02$, control $M = 1.36$, $SD = 0.78$; $z = 1.22$, $p = .11$) or low IQ (ASD $M = 0.75$, $SD = 0.76$, control $M = 1.00$, $SD = 1.05$; $z = 0.33$, $p = .37$).

Scene Description: Local measures

Table 7.11 presents the mean relative number of references made to surface details and the position of objects made by ASD and control participants. No significant group differences were found on either measure of local recall style, for the immediate (surface detail $z = 0.34$, $p = .37$; position $z = 0.68$, $p = .25$) or delayed condition (surface detail $z = 0.00$, $p = .50$; position $z = 0.23$, $p = .41$; all one-tailed tests). Furthermore, no group differences were found when comparing groups consisting of high IQ (all $z < 0.52$, $p > .60$), or low IQ participants (all $z < 1.49$, $p > .13$).

Table 7.11 Number of references to surface detail and position of objects as a proportion of the number of objects listed by individuals within each group overall, and split by high and low IQ: Mean (*SD*).

Group		Surface detail		Position	
		immediate	*delayed*	immediate	*delayed*
All	ASD	0.11 (0.16)	0.16 (0.27)	0.11 (0.20)	0.15 (0.23)
	Control	0.10 (0.11)	0.08 (0.11)	0.15 (0.21)	0.22 (0.36)
High IQ	ASD	0.13 (0.17)	0.19 (0.29)	0.13 (0.22)	0.18 (0.24)
	Control	0.13 (0.11)	0.10 (0.12)	0.16 (0.22)	0.25 (0.39)
Low IQ	ASD	0.04 (0.10)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	Control	0.00 (0.00)	0.00 (0.00)	0.10 (0.16)	0.08 (0.13)

The number of occasions that participants described objects that were not present in the underwater scene was very low. In the immediate description three participants from the control group (10%) described objects that had not been seen, while this occurred for only one participant in the ASD group (3%). Furthermore, five participants in the control group (17%) listed objects that were not present in the original scene in the delayed recall condition, while only one participant in the ASD (4%) gave this type of response. On all occasions the objects listed were related to an underwater scene. The difference in frequencies across groups did not reach significance (Fisher's exact test, immediate $p = .61$; delayed $p = .20$).

Correlations: Global measures

The global description composite showed no relation to age in the ASD group ($r_s = -.09$, $p = .65$) and the control group ($r_s = -.01$, $p = .95$). The global composite was, however, positively correlated with MA in the ASD group ($r_s = .44$, $p = .02$), which approached significance in the control group ($r_s = .34$, $p = .07$). A significant positive correlation was found between the global composite and VIQ and FIQ-BD in both groups ($r_s = .37$ to $.49$, all $p < .04$), but did not reach significance for FIQ (ASD $r_s = .34$, $p = .06$; control $r_s = .31$, $p = .09$), nor PIQ (ASD $r_s = .19$, $p = .31$; control $r_s = .15$, $p = .42$).

Correlations: Local measures

The relative number of references to surface details, averaged across immediate and delayed recall conditions, did not correlate with age or MA in either group (ASD $r_s < .20$, $p > .28$; control $r_s < -.25$, $p > .17$), or show any relationship to any IQ measure (ASD $r_s < .27$, $p > .15$; control $r_s < .18$, $p > .32$).

The relative number of references to position of objects, averaged across immediate and delayed recall conditions, showed a significant negative correlation with age in the ASD group (r_s = -.47, p = .009). This correlation was significantly different from that in the control group, where no reliable correlation held between references to object position and age (r_s = .15, p = .41; Fisher's z transformation, z_{r1-r2} = 2.45, p = .01). MA also correlated significantly with references to object position in the ASD group (r_s = .39, p = .03), which was not found in the control group (r_s = .20, p = .30; z_{r1-r2} = 0.78, p = .44).

The relative number of references to object position was strongly associated with all IQ measures in the ASD group (FIQ r_s = .62, FIQ-BD r_s = .65, VIQ r_s = .72, all p < .0005; PIQ r_s = .44, p = .02). No reliable correlation was observed, however, between this measure of local recall style and IQ in the control group (FIQ r_s = .02, FIQ-BD r_s = .05, PIQ r_s = -.08, VIQ r_s = .05, all p > .67; all z_{r1-r2} > 2.02, p < .04).

Recognition test

Table 7.12 presents the mean numbers (and percentages) of correct same/different judgments for each item condition in the recognition test for the ASD and control groups. One-sample t-tests were performed to verify that mean performance differed from chance. For both groups, detection rates for *same* and *different form* items were significantly above chance performance (all t > 3.74, p < .0005). The ability to detect *mirror image* items did not differ from chance in the ASD group ($t_{(28)}$= 0.99, p = .33), and was significantly below chance in the control group ($t_{(30)}$= -4.03, p < .0005).

Table 7.12 Number and *percentage* of correct same/different judgments for the three test item conditions in the Picture Memory Recognition Test for each group overall, and split by high and low IQ: Mean (*SD*).

	Group	N	Same (max = 10)	Different: Form (max = 10)	Different: Mirror Image[a] (max = 5)
All	ASD	29	6.38 (1.80) *64%*	6.62 (1.54) *66%*	2.28 (1.22) *46%*
	Control	31	6.94 (1.39) *69%*	6.03 (1.52) *60%*	1.68 (1.14) *34%*
High IQ	ASD	24	6.71 (1.63) *67%*	6.63 (1.47) *66%*	2.13 (1.26) *43%*
	Control	25	7.08 (1.35) *71%*	6.16 (1.49) *62%*	1.88 (1.13) *38%*
Low IQ	ASD	5	4.80 (1.92) *48%*	6.60 (2.07) *66%*	3.00 (0.71) *60%*
	Control	6	6.33 (1.51) *63%*	5.50 (1.64) *55%*	0.83 (0.75) *17%*

[a]ASD > Control, p = .03; ASD_{low} > $\text{Control}_{\text{low}}$, p = .004

Table 7.13 presents the ASD and control group means for the signal detection measure of sensitivity (A') for *different form* and *mirror image* test items. As the data failed tests of normality for both groups (all $p < .01$), non-parametric statistical analyses were used. The ASD group showed a greater bias towards making a "different" response compared to the control group overall, although the difference did not reach statistical significance (ASD $M = -.15$, $SD = .45$; control $M = -.39$, $SD = .35$; $z = 1.89$, $p = .06$). No group differences were found on A' for *different form* ($z = .09$, $p = .46$) and *mirror image* items ($z = 1.32$, $p = .09$). Furthermore, the high-functioning groups did not differ significantly in discriminating *different form* ($z = 0.03$, $p = .49$) or *mirror image* items ($z = 0.52$, $p = .30$). When comparing groups of low IQ, the ASD participants performed significantly better than control participants on A' for *mirror image* items ($z = 2.59$, $p = .005$) but not for *different form* items ($z = .09$, $p = .46$).

Recognition sensitivity was higher for *different form* items than *mirror image* items for the majority of participants in the ASD group ($z = 3.65$, $p < .0005$) and the control group ($z = 4.18$, $p < .0005$). This was also characteristic of the high IQ participants in the ASD group ($z = 3.43$, $p = .001$) and the control group ($z = 3.54$, $p < .0005$). Low IQ participants in the ASD group did not show a significant superiority for *different form* items over *mirror image* items ($z = 1.29$, $p = .20$) with only 3 of the 5 participants showing this pattern of performance. In comparison all 6 low-functioning control participants performed higher on the sensitivity measure for *different form* items relative to *mirror image* items.

Table 7.13 Recognition sensitivity (A') for detecting different form and mirror image test items on the Picture Memory Recognition Test for each group overall, and split by high and low IQ: Mean (*SD*).

	Group	N	A' Different Form	A' Mirror Image[a]
All	ASD	29	0.73 (0.11)	0.57 (0.21)
	Control	31	0.71 (0.14)	0.50 (0.22)
High IQ	ASD	24	0.74 (0.11)	0.57 (0.23)
	Control	25	0.73 (0.15)	0.55 (0.21)
Low IQ	ASD	5	0.65 (0.11)	0.57 (0.07)
	Control	6	0.65 (0.10)	0.31 (0.14)

[a] $ASD_{low} > Control_{low}$, $p = .005$

Correlations

Mean A' (averaged across *mirror image* and *different form* items) showed no relationship to age in the control group (r_s = -.02, p = .90) or ASD group (r_s = -.27, p = .17). The association between MA and mean A' approached significance in the control group (r_s = .35, p = .06), but showed no relationship in the ASD group (r_s = -.01, p = .96; z_{r1-r2} = 1.36, p = .18).

Recognition test performance was related to intellectual ability in the control group with significant correlations found between mean A' and FIQ (r_s = .41, p = .02), VIQ (r_s = .43, p = .02), and FIQ-BD (r_s = .35, p = .05), but not PIQ (r_s = .27, p = .15). In the ASD group, a significant association between mean A' and PIQ was found (r_s = .39, p = .04), but not for the remaining IQ variables (all r_s = .13 to .27, p > .15). The correlation coefficients did not differ significantly between groups (all z_{r1-r2} < 1.20, p > .22).

7.3.5 Summary and brief discussion of Picture Memory Task

The tendency to describe the underwater scene globally showed a clear developmental progression in the TD sample. No group differences were observed between the ASD and control groups in this response style, although the low IQ ASD participants showed a tendency not to use a global description at the outset. Local responses were defined as the number of references made to surface detail and object position relative to the number of objects listed. However, these two markers appeared to tap different abilities. The relative number of references to object position increased with age and IQ in the TD sample, particularly in male participants. No such age/IQ effects were observed in the relative number of references to surface detail. Males referred more to spatial position, while females described the surface detail of objects more often. These gender differences are in keeping with the spatial-verbal dichotomy often reported in gender research (McGee, 1979).

It was predicted from weak central coherence theory that individuals with ASD would tend to focus on surface features in the underwater scene and provide overly detailed descriptions. No differences were found, however, between the ASD and control groups in their style of recall using the two measures designed for the present study. Jolliffe and Baron-Cohen (2001b) also found no difference in the general style of descriptions made by high-functioning adults with ASD and ability-matched controls, although the ASD participants tended to not use bridging statements and misinterpret events occurring in a scene. Future work using a sequence of pictures that portray a narrative may be a more sensitive measure of coherence than a single complex scene as used in the present study.

Recognition memory for objects presented within a complex visual scene showed a moderate increase with age and intellectual ability in the TD sample. As predicted, detecting changes in surface form (i.e., left/right orientation) was significantly more difficult than detecting changes in semantic form, with performance levels at or below chance for the majority of TD participants. The ASD group also did not differ from chance in detecting change in orientation, but performed significantly higher than the control group who performed below chance levels. ASD participants with low IQ were superior to their matched control participants at correctly detecting change in orientation. The prediction from weak central coherence theory that the majority of individuals with ASD would show greater proficiency on the recognition test overall was therefore not fully supported by the present findings.

Although the attempt was made to create an open-ended task, it was noted that some participants were aware that they might be tested on remembering details of the picture. This level of expectation, particularly in high-functioning individuals, may have resulted in differences based on cognitive ability, rather than style.

Two continuous measures of central coherence processing style were taken from the Picture Memory Task. From the recognition test, mean A' was taken as an index of the ability to discriminate correctly both form and surface changes, with high scores indicating weak central coherence. The 'global composite' constructed from the verbal descriptions, which incorporated all four indices of the tendency to describe the underwater scene in a wholistic manner, was taken as the second continuous measure of central coherence, with high scores indicating strong coherence. Participants were categorised as showing weak central coherence if their mean A' index was greater than 0.75. Weak central coherence was also indicated categorically if the individual failed to provide a global description of the underwater scene in both the immediate or delayed recall conditions.

7.4 Fragmented Pictures Task

The ability to make sense out of incomplete, fragmented visual stimuli is an important perceptual skill and tasks tapping this ability were among the earliest neuropsychological measures developed (Lezak, 1995). Gestalt completion or perceptual closure tasks, such as the Street Completion Test (Street, 1931) and the Gestalt Closure subtest from the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983), require the participant to identify partially completed drawings of common objects. Success relies on strong visual coherence in order to "combine disconnected, vague, visual stimuli into a meaningful whole" (Carroll, 1993, p. 308). As such, a fragmented picture task was considered a potential measure of configural processing at a visual-semantic level.

Gollin (1960) devised a measure of gestalt completion where objects are represented by a series of five drawings that vary in degree of fragmentation. The most fragmented image is presented first and is replaced by a more complete image until naming is achieved or the complete picture is presented. In contrast to naming a single image, this method is applicable to a wide range of participants and allows greater individual variation in performance. Gollin administered the task to preschool-aged children and found a significant correlation between chronological age and the amount of image required for recognition. He also found that adults recognised objects at a greater degree of fragmentation than did children. With normal aging, this ability declines and elderly participants have been found to require more of the image for identification than young adults (Patterson, Mack, & Schnell, 1999).

Snodgrass, Smith, Feenan, and Corwin (1987) developed the Fragmented Picture-Completion task based on the Gollin Figures (Gollin, 1960). Line drawings of common objects, taken from the Snodgrass and Vanderwart (1980) picture set, were fragmented by the random deletion of pixels into a series of eight images (see Figure 7.11 for an example). The most fragmented level is presented on a computer screen and participants can advance through each successive image until they recognise the complete form. Although the Fragmented Picture-Completion task was devised for experiments on implicit memory, a wealth of normative data exists on identification thresholds for 150 picture stimuli (Koch, Abbey, & Schmidt, 1995; Snodgrass & Corwin, 1988; Wyatt, Conners, & Carr, 1998). The present study used the images developed by Snodgrass et al. in an adaptation of their Fragmented Picture-Completion task.

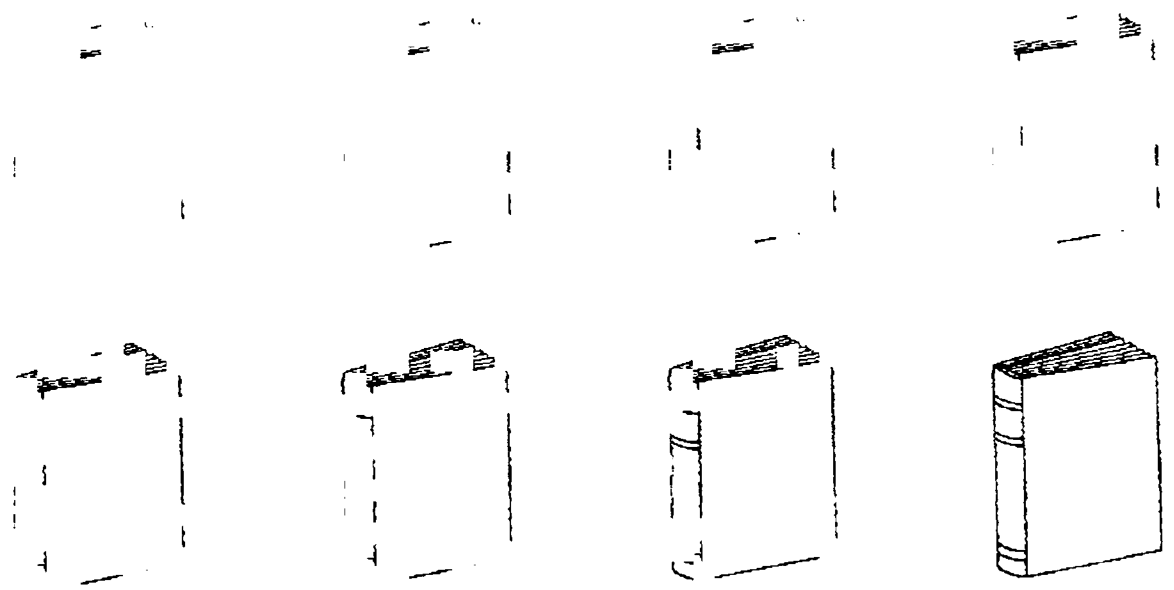


Figure 7.11 The eight levels of fragmentation for the book stimulus from the Fragmented Picture-Completion task (Snodgrass et al., 1987).

Lezak (1995) reported small sex differences favouring males have been found on standard gestalt completion tasks, which use single picture stimuli. To the author's knowledge, no investigation of gender effects on the Gollin Figures (Gollin, 1960) or the Fragmented Picture-Completion Task (Snodgrass et al., 1987) have been reported in the literature.

Picture integration in ASD

Visual integration abilities have been explored in ASD to a limited extent. Jolliffe and Baron-Cohen (2001a) investigated this ability as a means to test weak central coherence theory in adults with either autism or Asperger syndrome. They found that individuals with an ASD were less able to integrate visual elements using a modified version of the Hooper Visual Organisation Test (Hooper, 1983) where objects have to be identified from parts positioned randomly on a page. Interestingly, Jolliffe and Baron-Cohen found no impairment in their participants with ASD to identify an object from a single piece, but a specific deficit in the ability to conceptually integrate single elements in order to form a meaningful whole (when potentially identifying single details were absent).

Although Mottron and Belleville (1993) did not set out to assess weak central coherence, they did assess the visual integration capabilities of their single case of a savant artist with Asperger syndrome. No differences in performance to control participants were found however in the ability to recognise degraded pictures using Gollin's graded picture series (Gollin, 1960). Furthermore, no impairment was found on the Hooper Visual Organisation Test. Given that this was a single case study of an artistically talented individual, a deficit in visual integration may still be a valid marker of weak central coherence in individuals with ASD and provide the means to explore individual differences in central coherence.

Predictions

The ability to cohere fragmented elements into a meaningful whole and quickly identify an object is predicted to increase with development in TD individuals. Individual differences in visual integration ability are also predicted, over and above age and IQ effects. Weak central coherence theory predicts that individuals with ASD will be less proficient at integrating the degraded information so as to identify the whole object. Individuals with ASD are therefore expected to require significantly more of the image for correct identification, relative to age- and IQ-matched control participants.

7.4.2 Method

Materials

Ten picture sequences were selected from the Fragmented Picture-Completion task (Snodgrass et al., 1987; see Appendix M). It was ensured that pictures could not be identified on the basis of individual parts alone, such that critical features (e.g., an eye or a tail in a picture of an animal) were not present at the most fragmented levels. This procedure was adopted by Jolliffe and Baron-Cohen (2001a) in their modification of the Hooper Visual Organisation Test, and ensures that the task is assessing visual integration ability, rather than successful recognition based on an isolated element. Pictures were also selected on the basis of high name agreement and familiarity as rated by young children (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997).

The Fragmented Pictures Task was presented using SuperLab Pro software controlled by a laptop computer. Pictures appeared within a 3.25 by 3.25 inch square, positioned centrally, on a 15-inch computer touch screen.

Procedure

To introduce the concept of fragmentation a complete picture of a chair first appeared on the computer screen. Participants were told to watch the picture as it was slowly going to disappear. The image was successively replaced every five seconds by a less complete image and the participant was encouraged to say when they could no longer recognise the picture as a chair. It was then explained that on this task they would see the opposite; pictures of objects were slowly going to appear on the screen. Their task was to watch the screen and tell the researcher as soon as they could recognise the picture.

Each fragmented image was presented one at a time for five-second exposures, from the most fragmented image (first frame) through to the complete image (eighth frame). When the participant gave a response the researcher immediately suspended the program (including the timer) and a 3-point rating scale would appear on the screen, replacing the fragmented picture sequence. The rating scale was numbered 1, 2, 3, with the two ends labelled *"not that sure"* (1) and *very sure"* (3). Participants were asked to use the scale to indicate how certain they were of their answer by either pointing to the corresponding number, or saying the number/label aloud[8]. The researcher wrote down the rating and then informed the participant of the correctness of their response. If they were incorrect, they were told to keep looking and the program (including the timer) restarted from the

[8] This rating scale was included to check for a possible confound between relative inability to cohere the figure parts and a relative unwillingness to guess. It was hypothesised that the ASD group may be more reluctance to guess and therefore withhold responding until they were absolutely sure of their answer. However, as the ASD group did not show any hesitancy in providing guesses, the findings from the certainty ratings are not presented here.

beginning of the frame at which it was suspended. If they were correct, the researcher congratulated the participant and moved the program on manually until the complete image appeared. The researcher then began a new trial starting with the most fragmented image of the next object.

Time was recorded from the presentation of the most fragmented image until the participant provided a correct response; this included a summation of times when incorrect responses were given. The frame number (ranging from one to eight) at which each item was correctly identified and the number of incorrect responses were also recorded for each participant.

Two sets of 14 picture sequences were first piloted with 44 TD children (7 to 15 years). The first set consisted of pictures of moderate difficultly (i.e., were correctly identified by 35 percent of adults by the fourth frame), and the second set contained pictures of high difficultly (i.e., were correctly identified by 35 percent of adults by the sixth frame) according to the norms collected by Snodgrass and Corwin (1988). The final set of 10 picture sequences were selected from items of moderate difficultly. Performance on these items produced good inter-participant variability, with no suggestion of ceiling effects at the age levels tested. Items of moderate difficulty were also considered to be more applicable for participants with a range of ability levels, especially those with suspected global processing deficits. Items were also discarded when two or more participants in the pilot sample were unable to name the picture at the final frame.

7.4.3 Typical Development Results

All TD participants were administered the Fragmented Pictures Task and age group results are presented in Table 7.14. Identification rates for each picture were at ceiling except for the *pear* item, which three young participants (9 to 10 years) failed to name in its complete form. One 22-year-old female (FIQ 97, VIQ 108, PIQ 86) was deemed an outlier in her age group (mean response time 32.1 seconds, mean frame number for correct identification 6.90). Her data were removed from analysis, which resulted in all age group data approximating a normal distribution with no significant skewness. The number of incorrect guesses was heavily positively skewed (z-scores ranged from 3.5 to 5.0), while the mean number of items correctly identified at the 6th, 7th or 8th frame was not normally distributed, so non-parametric statistical analyses were used for these two variables.

A two-way ANOVA performed on the mean frame number for correct detection showed a significant main effect of age group ($F_{(3,195)} = 12.29$, $p < .0005$). No significant main effect of gender was found ($F_{(1,195)} = 0.81$, $p = .37$), nor interaction between age group and gender ($F_{(3,195)} = 0.57$, $p = .62$). As shown in Figure 7.12, male and female participants performed at a similar level within each age group. Adjusting for the effects of FIQ did not alter the pattern of findings. Post hoc analyses (Tukey's HSD) revealed that the age group effect was due to the youngest age group requiring more frames before correct detection than all older age groups (all $p < .004$), while the three older age groups did not differ from each other. Using mean response time for correct identification produced the same pattern of results, although participants in the 11-13 year group were found to take significantly longer to identify each picture than the oldest age group ($p = .02$).

Table 7.14 Fragmented Pictures Task results by age group: Mean (*SD*).

Age group (years)	Mean frame number for correct identification (max = 8)[a]	Mean response time for correct identification (seconds)[b]	Total number of incorrect guesses[c]	Total number of items identified at 6th, 7th or 8th frame (max = 10)[d]
8 - 10	5.41 (0.44)	25.4 (2.0)	3.00 (3.02)	4.06 (1.37)
11 - 13	5.07 (0.51)	23.2 (2.5)	2.21 (1.78)	3.16 (1.62)
14 - 16	5.01 (0.49)	22.7 (2.6)	1.43 (1.54)	2.90 (1.69)
17 - 25	4.84 (0.46)	21.9 (2.2)	1.36 (1.32)	2.25 (1.65)

[a] 8-10 < 11-13, 14-16, 17-25, $p < .004$
[b] 8-10 < 11-13, 14-16, 17-25, $p < .0005$, 11-13 < 17-25, $p = .02$
[c] 8-10, 11-13 > 14-16, 17-25 $p < .01$
[d] 8-10 > 11-13, 14-16, 17-25 $p < .006$, 11-13, 14-16 > 17-25, $p < .05$

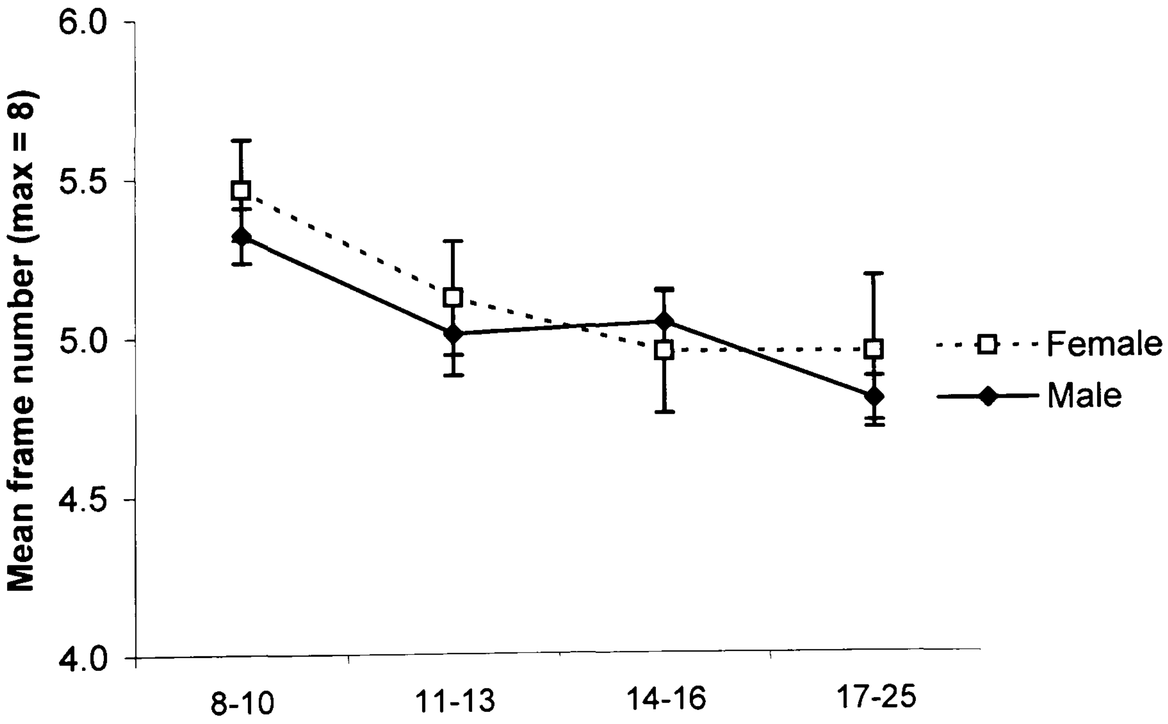


Figure 7.12 Mean frame number for correct identification on the Fragmented Pictures Task for male and female participants in each age group. Error bars show 95% confidence intervals.

The mean number of incorrect guesses showed a significant effect of age group ($\chi^2 = 14.88$, $p = .002$), with the two younger age groups making more guesses than the two older age groups (all $z > 2.57$, $p < .01$). No gender effects were found on the total number of incorrect guesses, over the whole sample ($z = 1.21$, $p = .23$), and within each age group (all $z < 1.31$, $p < .19$).

As a final measure of task performance, participants were compared on the number of pictures that were recognised only near their most complete form (i.e., either at the 6^{th}, 7^{th} or 8^{th} frame; see Table 7.14). Again a significant effect of age group was found on this measure ($\chi^2 = 36.25$, $p < .0005$) with pictures identified at an earlier stage in older participants. No gender effects were found on this variable (all $z < 1.61$, $p > .11$).

Correlations

Across the TD sample there was a significant correlation between performance on the Fragmented Pictures Task and age, with older participants correctly identifying objects in faster time ($r = -.46$) and after fewer frames ($r = -.38$, all $p < .0005$). This relationship remained significant within both male and female participants (all $r > -.32$, $p < .002$).

The only IQ measure to correlate significantly with task performance was PIQ (mean frame number for correct identification, $r = -.16$, $p = .02$; response time, $r = -.21$, $p = .003$). The correlation coefficients fell below significance for the remaining IQ variables (all $r < -.10$, $p > .14$). When the gender groups' performances were considered separately, PIQ and task performance correlated significantly in males (mean frame number, $r = -.29$, $p = .004$; mean response time, $r = -.32$, $p = .001$), but not in females (mean frame number, $r = -.02$, $p = .87$, $z_{r1-r2} = 1.96$, $p = .05$; mean response time, $r = -.07$, $p = .46$, $z_{r1-r2} = 1.84$, $p = .06$).

7.4.4 ASD and Control Results

One ASD participant was not administered the Fragmented Pictures Task, although the groups remained well matched on age and IQ measures after removing this participant from analyses. Accuracy rates were very high for both groups and only two ASD participants (aged 9 and 16 years) failed to name a picture in its complete form (*pear*, *television*).

Table 7.15 presents descriptive statistics for the Fragmented Pictures Task, split by group. As the number of incorrect guesses was positively skewed in each group (ASD $z = 2.30$, control $z = 5.74$), non-parametric analyses were used for this variable. The mean frame number and response time for correct identification, and mean number of items identified at the 6^{th}, 7^{th} or 8^{th} frame were not significantly skewed and passed tests of normality. Parametric statistical analyses were used for these variables.

Independent t-tests (one-tailed) showed a trend towards a significant group difference on mean frame number for correct detection ($t_{(59)} = 1.58$, $p = .06$), and a significant difference on mean response time for correct detection ($t_{(59)} = 2.12$, $p = .02$). The ASD group took longer to recognise each object from its separate parts, suggesting some difficulty in cohering the fragmented information.

When including participants with high IQ only, significant group differences were not found for the mean number of frames for correct detection ($t_{(47)} = 0.45$, $p = .33$), or for the mean response time per item ($t_{(47)} = 1.34$, $p = .09$). Group differences were found, however, between groups of low IQ, with the ASD group requiring significantly more time before correctly identifying each object ($z = 1.92$, $p = .03$), and doing so at a higher frame number ($z = 2.33$, $p = .01$).

Table 7.15 Fragmented Pictures Task results by group overall, and split by high and low IQ: Mean (*SD*).

	Group	Mean frame number for correct identification (max = 8)[a]	Mean response time for correct identification (seconds)[b]	Total number of incorrect guesses[c]	Total number of items identified at 6th, 7th or 8th frame (max = 10)[d]
All	ASD	5.32 (0.51)	24.7 (2.6)	2.87 (2.58)	4.13 (1.63)
	Control	5.13 (0.45)	23.4 (2.3)	2.00 (2.68)	3.13 (1.80)
High IQ	ASD	5.19 (0.47)	24.1 (2.3)	2.75 (2.21)	3.71 (1.52)
	Control	5.13 (0.46)	23.2 (2.3)	1.48 (1.96)	3.24 (1.85)
Low IQ	ASD	5.83 (0.30)	27.3 (2.1)	3.33 (3.98)	5.83 (0.75)
	Control	5.10 (0.48)	24.2 (2.5)	4.17 (4.22)	2.67 (1.63)

[a] $ASD_{low} > Control_{low}$, $p = .01$
[b] ASD > Control, $p = .02$, $ASD_{low} > Control_{low}$, $p = .03$
[c] $ASD_{high} > Control_{high}$, $p = .02$
[d] ASD > Control, $p = .01$, $ASD_{low} > Control_{low}$, $p = .003$

As response time was a summation of all frame exposures including incorrect responses (i.e., the timer is stopped as soon as a response is made, and resumes from the beginning of the last frame if an incorrect response is made), longer detection times may directly relate to a greater number of incorrect guesses. The non-parametric Mann-Whitney *U* test did not, however, detect group differences in the number of incorrect guesses for the whole sample ($z = 1.71$, $p = .09$, two-tailed). ASD participants with low IQ made fewer guesses than their matched control group, although not significantly ($z = 0.49$, $p = .63$). ASD participants with high IQ did, however, make significantly more incorrect guesses than their control group ($z = 2.34, p = .02$).

The number of pictures correctly identified at the 6th, 7th, or 8th frame was significantly higher in the ASD group than the control group ($t_{(59)} = 2.28$, $p = .01$, one-tailed). No group difference was found on this variable for groups of high IQ participants ($t_{(47)} = 0.97$, $p = .17$), but was found between groups of low IQ ($z = 2.77, p = .003$).

A qualitative analysis of the nature of incorrect guesses was conducted for ASD and control participants. There was no indication of 'isolate' responses (Hooper, 1983) from either group; that is, few responses could be interpreted as being based on a local detail.

Correlations

No significant correlation was found between age and task performance in the ASD group (mean frame number $r = .06$, $p = .75$; mean response time $r = -.02$, $p = .91$) or the control group (mean frame number $r = -.15, p = .41$; mean response time $r = -.19, p = .30$). MA, in contrast, showed a significant correlation with task performance in the ASD group (mean frame number $r = -.44$, $p = .02$; mean response time $r = -.47$, $p = .01$), but did not reach significance in the control group (mean frame number $r = -.11$, $p = .56$; mean response time $r = -.30, p = .11$; all $z_{r1\text{-}r2} < 1.34, p > .18$).

PIQ showed a weak association with mean response time for correct identification in the control group ($r = -.32$, $p = .08$), while all other IQ variables showed no relationship with task performance ($r = -.20$ to $-.08$, all $p > .10$). A strong association between IQ and visual integration ability was found in the ASD group however, with high IQ relating to the correct identification of objects at an earlier level of fragmentation ($r = -.48$ to $-.44$, all $p < .02$). The magnitude of the correlation coefficients between the mean frame number and FIQ, VIQ, and FIQ-BD, were significantly higher in the ASD group than the control group (all $z_{r1\text{-}r2} > 1.98, p < .05$).

7.4.5 Summary of Fragmented Pictures Task

Visual integration ability, as assessed by the Fragmented Pictures Task, showed a significant improvement with age in the TD sample, particularly after age 10-years. The ability to cohere fragmented pieces of information was associated with visuo-spatial ability in the TD sample, and specifically males, but not with verbal ability or overall intelligence. No effect of gender was found in level of performance on the Fragmented Pictures Task.

As predicted by central coherence theory, individuals with ASD were significantly impaired in their ability to integrate fragments of information in order to identify single objects. This effect was stronger in low-functioning individuals with ASD and consequently, a strong association was found between visual integration ability and overall intelligence in the ASD group, which was not apparent in the control group. The current findings support those of Jolliffe and Baron-Cohen (2001a).

As mean response time for correct identification of fragmented objects was a sensitive measure of individual differences and discriminated between ASD and control groups, this variable was used as the main index of processing style for the Fragmented Pictures Task, with longer times indicating weak coherence. Participants were classified as showing weak coherence if they correctly identified five or more pictures (out of a total of ten) in the final frames (i.e., at the 6th, 7th, or 8th frame).

7.5 General Discussion

A summary of experimental findings from the high-level visuo-spatial coherence tasks is presented in Table 7.16 for the TD sample and in Table 7.17 for the ASD and control groups (see Appendix Q for effect sizes). Table 7.18 presents the percentage of participants in each group categorised as showing weak central coherence on each visuo-spatial measure.

Local and global processing of high-level visuo-spatial stimuli was assessed through various means: drawing style, benefit from meaning in recall, recognition memory for visual details, verbal description of a detailed scene, and the ability to integrate separate visual elements. With the exception of benefit from meaning in recall accuracy, developmental effects were found in the TD sample with older participants tending to use a more global drawing style, provide more global descriptions of the scene, and show greater proficiency in integrating degraded visual information. Recognition memory for visual details improved with age and intellectual ability, suggesting a developmental progression from global to local processing when a local bias was advantageous to task demands. Despite the common finding of gender effects on visuo-spatial tasks, male and female participants tended to perform equivalently on the coherence measures. The exception was in drawing style, where males showed a local bias and tended to draw in a more piecemeal and fractionated manner than females. Despite the evidence of developmental effects on task performance, there was wide variability between TD participants and a significant proportion of the variance from each dependent variable could not be explained by age, IQ, and gender (see Table 7.16), and may reflect differences in cognitive style.

Contrary to predictions from weak central coherence theory, a local processing bias was not pervasive in individuals with ASD across the majority of high-level visuo-spatial tasks. There was no suggestion that individuals with ASD used a local drawing style or showed less benefit from meaning in the recall of geometric figures than control participants. Individuals with ASD did not differ from their matched controls in the way they described a complex scene, although ASD participants with low IQ tended not to give a global description initially. Recognition memory for local detail changes was higher in the ASD group than the control group (particularly in participants with low IQ), however the use of signal detection procedures that took response bias into account did not show the groups to differ significantly.

Results from the Fragmented Pictures Task did support predictions from weak central coherence theory, and suggest that individuals with ASD had difficulty with the integration of fragmented visual information. In contrast to the other high-level visuo-spatial coherence tasks described, the Fragmented Pictures Task did not place demands on production or memory processes, other than object recognition and naming. This may have resulted in a 'purer' measure of central coherence, not contaminated by other high-level processes required for successful task performance.

The drawing and scene description tasks were designed to be as open-ended as possible, in order to capture the natural processing style of participants in a non-directive way. In the Picture Memory Task especially, no prior warning was given as to the nature of the recall and recognition testing (unlike Story Memory, Chapter 8). However, participants may have tried to predict the task requirements (e.g., verbally count/label single objects within the scene), and hence strategic or ability factors rather than style may have been tapped. As discussed in Section 7.2.7, the structured approach taken to assess coherence in drawing style may not have been able to detect group differences. A more subjective measure will be considered for future work on drawing style.

The coherence indices constructed from the high-level visuo-spatial tasks are revisited in Chapter 9, which explores the inter-correlations of the measures, as well as associations between measures from the other processing domain/level.

Table 7.16 Summary of high-level visuo-spatial domain findings in the TD sample.

Visuo-spatial Task	Dependent Variable (Direction of high values)	Age effects?	IQ effects?	Gender effects?	Proportion of variance not explained by age, IQ and gender	Interactions
Drawing Task	Coherence index Rey/Pram copy (Strong CC)	Yes, positive	Yes, positive (FIQ, PIQ)	F > M	78%	
	Effect of meaning on recall accuracy (Strong CC)	No	No	No	99%	Positive correlation with VIQ in M only
Picture Memory	Global description composite (Strong CC)	Yes, positive	Yes, positive	No	86%	Positive correlation with PIQ in F; all IQ correlations in M
	Mean A' (Weak CC)	Yes, positive	Yes, positive	No	95%	
Fragmented Pictures Task	Mean RT for correct identification (Weak CC)	Yes, negative	Yes, negative with PIQ only	No	80%	Negative correlation with PIQ in M only

Note: M = males, F = females

Table 7.17 Summary of high-level visuo-spatial domain findings in the ASD and control groups.

Visuo-spatial Task	Dependent Variable (Direction of high values)	Group effects?	Age effects?	MA effects?	IQ effects?	Other comments
Drawing Task	Coherence index Rey/Pram copy (Strong CC)	No	No	No (positive, n.s.)	No (positive, n.s.)	
	Effect of meaning on recall accuracy (Strong CC)	No	No	No	Positive with PIQ in ASD only	ASD_{high} < $Control_{high}$ (not as predicted)
Picture Memory	Global description composite (Strong CC)	No	No	Yes, positive	Yes, positive (not PIQ)	
	Mean A′ (Weak CC)	No	No	Positive in Control only	Positive FIQ/VIQ in Control; positive PIQ in ASD	ASD_{low} > $Control_{low}$
Fragmented Pictures Task	Mean RT for correct identification (Weak CC)	ASD > Control	No	Yes, negative	Negative in ASD only	ASD_{low} > $Control_{low}$

Table 7.18 Percentage (and N/total N) of participants in each group who showed weak central coherence on each task.

Visuo-spatial Task	Marker for weak CC	8-10	11-13	14-16	17-25	ASD	Control
Drawing Task	More than one zero rating of style Rey & Pram copy[a]	32% (17/54)	26% (11/43)	35% (18/51)	13% (7/56)	23% (7/31)	26% (8/31)
	Construction order < 2.15 (max = 4) (tendency to draw local features in first third of drawing)	43% (23/54)	30% (13/43)	29% (15/51)	20% (11/56)	32% (10/31)	42% (13/31)
Picture Memory	A' > .75	17% (9/54)	26% (11/43)	33% (17/51)	38% (21/56)	24% (7/29)	23% (7/31)
	Never provided a global description immediate/delayed recall[b]	35% (19/54)	23% (10/43)	16% (8/51)	5% (3/56)	23% (7/30)	23% (7/31)
Fragmented Pictures Task	5 or more pictures (max 10) identified at 6th, 7th or 8th frame[c]	41% (22/54)	19% (8/43)	20% (10/51)	6% (3/55)	43% (13/30)	26% (8/31)

[a] 8-10, 14-16 > 17-25, $p < .04$
[b] 8-10, 11-13 > 17-25, $p < .04$, 8-10 > 14-16, $p = .05$
[c] 8-10 > 11-13, 17-25, $p < .04$, 14-16 > 17-25, $p = .04$; ASD > Control, $p = .08$

Chapter 8. High-level verbal coherence tasks

8.1 Introduction

Chapter 8 presents experimental findings from three tasks designed to assess the contextual aspects of language processing. Individual differences in local/global processing of verbal material was assessed by the tendency to read for meaning, the degree of verbatim recall and recognition for short stories, and the nature of completions given to sentence stems. The prediction from central coherence theory that individuals with ASD will be impaired in their ability to cohere verbal information and show preference for surface-level over semantic-level information was tested. As all tasks involved use of the English language, data were compared in the TD sample with the inclusion and exclusion of participants with ESL.

8.2 Sentence Completion Task

In the Sentence Completion Task participants are asked to complete sentence stems such as *'You can go hunting with a knife and…'*. Globally meaningful completions such as *'gun'* or *'catch a bear'* show intact or strong coherence, while local completions such as *'fork'* suggest weakened coherence or a tendency to prefer local over global coherence. This task was first described in a study of the broader autism phenotype (Happé et al., 2001), which showed that the tendency to make globally-inappropriate local sentence completions was characteristic of the parents, and especially the fathers, of boys with ASD. Although the absolute number of local completions made by participants was few, individual variation was found when long hesitations were also included as an index of test performance. The Sentence Completion Task was therefore included in the present study as a measure of coherence in the verbal domain.

Sentence Completion in ASD

Evidence has emerged from the literature pertaining to coherence in autism that individuals with ASD have difficulty processing linguistic information for meaning, especially connected text (see Chapter 2, Section 2.4.3). For example, reduced use of sentence context for disambiguation of homographs is now well replicated (Frith & Snowling, 1983; Happé, 1997; Lopez & Leekam, 2003). An impressive set of studies by Jolliffe and Baron-Cohen has shown poor integration of verbal material (e.g., choosing bridging coherent sentence, recognising rare context-dependent sentence meanings) even in individuals with Asperger syndrome (Jolliffe & Baron-Cohen, 1999, 2000). Of relevance to the present study is the work by Frith and Snowling (1983) who demonstrated difficulties processing words in context in ASD, across a range of tasks including selecting the right word to fill a gap in a story text. Children with autism were more likely to insert a word that was semantically inappropriate for the context (although often syntactically correct) compared to control children matched on single-word reading ability.

Sentence Completion in Typical Development

Sentence completion tasks are commonly used as projective or personality assessment instruments (Holaday, Smith, & Sherry, 2000). They have proved to be a successful technique for use with children, providing an indirect means to study the expression of personality traits such as aggression (Beech & Graham, 1967), as well as a tool to facilitate verbal responses from severely disturbed and disabled children (Cobrinik, 1977; Sudhalter, Maranion, & Brooks, 1992). Sentence completion is also required in certain executive function tasks, notably the Hayling Sentence Completion Test (Burgess & Shallice, 1996). The emphasis in this task is on a contrast between a straightforward meaningful completion (thought to be effortless), and deliberately meaningless completion (effortful inhibition) conditions. The ability to inhibit a meaningful response in such tasks has been found to increase with development (Lorsbach & Reimer, 1997).

Although completion norms for standard sentence stems have been collected, these are primarily from adult populations (e.g., Bloom & Fischler, 1980). Few studies have examined the extent to which the meaningfulness of sentence completions varies with development, specifically the degree of local and global coherence. In an early study, Beech and Graham (1967) found that a sentence completion task could differentiate aggressive from non-aggressive children, but task performance was not significantly influenced by gender, age, or intelligence in TD children aged 12 to 15 years. Cobrinik (1977) used a sentence completion task to study language abilities of children (12 to 17 years) with a range of psychiatric disorders, including schizophrenia and mental retardation. An analysis of error types revealed that the most common errors could be classified as 'local completions', where responses were made in relation to the immediate cue in isolation from its general context. This occurred more often when the immediate cue contained an emotional connotation (e.g., *'Now dry your eyes and stop … hurting him'*). The frequency of error types was not gathered from TD children in this study however. The present study set out to provide normative data on the frequency of local, globally-inappropriate, completions to sentence stems in a large TD sample, to establish the extent to which age, intelligence and gender may influence the nature of local/global sentence completions.

Predictions

According to weak coherence theory, individuals with ASD may not readily process linguistic information for global meaning, and instead show a piecemeal processing style.

As a consequence they are suggested to provide completions to sentence stems that are meaningful at a local level, but not at an overall global level. Thus, the prediction was made that individuals with ASD would provide locally determined endings to sentences more often than control participants. As the frequency of local completions to sentence stems has not been studied extensively in the TD literature, no specific predictions of age, IQ and gender effects to task performance were made.

8.2.2 *Method*

Materials

The Sentence Completion Task, as described in Happé et al. (2001), consisted of ten sentence stems, designed to produce a conflict between making an appropriate global completion and succumbing to a locally cued associative response. For example, a local response to the sentence *"The fireman carried the bucket and ..."* would be *"spade"*; this response completes the final two words in isolation, but is meaningless in the context of the whole sentence. As a modification to the original task, five new sentence stems were included that did not produce a local-global conflict for their completion (e.g., *'I was given a pen and...'*). The filler sentences were added in order to mask the nature of the test items and make the task more naturalistic. These control items were interspersed with test items in a set order for all participants (see Appendix N for the complete set of sentence stems).

Procedure

The experimenter told the participant that she was going to read some sentences aloud, and that his/her task was to *"Say something to finish off the sentence"*. The task began with a practice control sentence, that was not designed to elicit a locally cued response: *'He cleaned up the mess with a brush and ...'*. Completions could be a single word or phrase, although the experimenter did not make reference to this unless the participant directly asked. The experimenter would try to prevent the participant repeating the entire sentence in their response and only provide a completion. Responses were written down by the experimenter and audio taped for later scoring. The time it took for the participant to make a response from the end of the each sentence stem was recorded to the nearest 0.5 second. The maximum time allowed for each completion was 20 seconds.

8.2.3 *Scoring*

(1) *Error Score.* Scoring was based on participants' first response. A 3-point scoring system was developed to capture the range of responses that were produced for the ten conflict sentence stems: 0-points was assigned for a globally meaningful completion that was produced within 10 seconds; 1-point was assigned when no response was given, the response delay was greater than 10 seconds, or the response was an "odd" completion to

the sentence, but not an obvious local completion (e.g., a repetition, or local associate to another word in the sentence); 2-points were assigned for a response that was a local associate to the end of the sentence, and not meaningful in the context of the whole sentence. An example of a local response to the sentence "*The sea tastes of salt and …*" would be "*pepper*", whereas "*water*" would not be scored as a local error (even though "salt and water" might be considered associates) since this response is appropriate to the meaning of the whole sentence. Inter-rater agreement from a second coder, rating responses from a third of the TD, ASD, and control samples, was greater than 95%.

(2) *Number of Local Responses.* The total number of local completions for each participant (i.e., number of 2-point responses) was used as a measure of extreme local performance.

(3) *Response time.* The mean response time for the ten conflict items for each participant was also used a measure of processing time to produce a completion.

8.2.4 Typical Development Results

Although all TD participants received the Sentence Completion Task, the following analysis is based on participants who had English as their first language (N = 192). It was considered most appropriate to include only participants whose first language would likely contain classic English phrases, such as *salt and pepper*, and *fish and chips*. The IQ and gender distribution overall and within each age group did not differ from the complete sample after the exclusion of the 12 participants with ESL (see Appendix D).

Table 8.1 presents the results from the Sentence Completion Task for each age group. Analysis of estimates of skewness and kurtosis showed that assumptions of normality were violated for the error score, number of local completions, and mean response time data in each age group (z-scores ranged from 3.0 to 7.9 for skewness and from 0.4 to 15.9 for kurtosis). Non-parametric statistics were therefore used and all outliers in the data were retained.

Table 8.1 Sentence Completion Task results by age group: Mean (*SD*).

Age group (years)	*N*	Error score[a]	Number of local completions	Response time per completion (seconds)[b]
8-10	54	2.57 (2.18)	0.74 (1.07)	3.93 (1.82)
11-13	40	2.65 (2.58)	0.93 (1.27)	3.12 (1.72)
14-16	49	1.51 (2.12)	0.57 (1.02)	2.19 (1.39)
17-25	49	1.45 (1.60)	0.45 (0.68)	2.48 (1.56)

[a]8-10, 11-13 > 14-16, 17-25, $p < .03$
[b]8-10 > 11-13, 14-16, 17-25, $p < .02$; 11-13 > 14-16, 17-25, $p < .03$

Although no age group effects were found on the mean number of local completions made ($\chi^2 = 4.66$, $p = .20$), the error score showed a significant effect of age group ($\chi^2 = 15.7$, $p = .001$). The two youngest age groups scored significantly higher on this measure than the two oldest age groups (all $z > 2.24$, $p < .03$). The mean time for participants to provide a completion also showed a significant effect of age group ($\chi^2 = 42.77$, $p < .0005$). The youngest age group took significantly longer to provide a response than all older age groups (all $z > 2.61$, $p < .02$). The 11-13 year group was also slower to provide a response than the two older age groups (all $z > 2.18$, $p < .03$).

Male and females overall, and within each age group, did not differ statistically on any of the three measures from the Sentence Completion Task (all $z < 1.83$, $p > .06$). However, an effect of age group on the error score was only significant in male participants ($\chi^2 = 12.32$, $p = .01$), and not female participants ($\chi^2 = 5.35$, $p = .15$). Males in the two youngest age groups scored significantly higher on the error score than the two oldest age groups (all $z > 2.25$, $p < .03$). The greater effect of age group for male participants than for female participants on the error score is depicted in Figure 8.1.

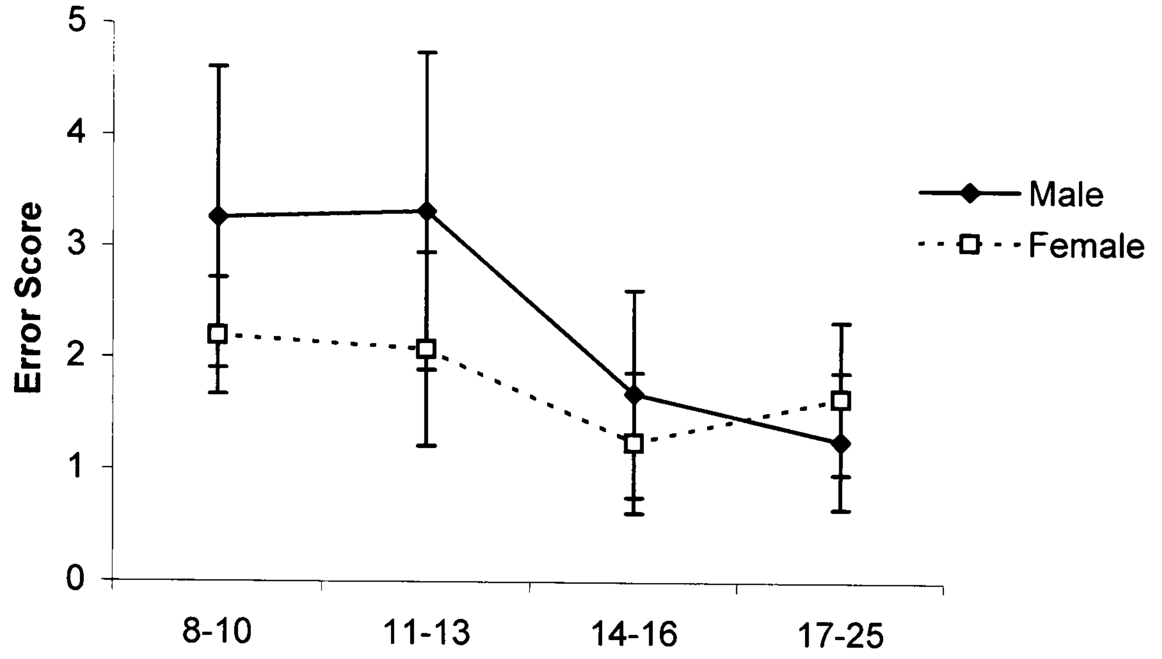


Figure 8.1 Mean error score on the Sentence Completion Task for male and female participants for each age group. Error bars show 95% confidence intervals.

Correlations

Speed-accuracy trade-off effects were examined in the TD sample. Local completions were associated with fast responses in female participants, with a significant negative correlation found between mean response time and frequency of local responses (r_s = -.26, p = .01). No such association was found in male participants however (r_s = .001, p = .99; z_{r1-r2} = 1.83, p = .06).

As the error score was the most continuous index of coherence from the Sentence Completion Task, this variable was used in correlational analyses. A significant negative correlation between age and sentence completion performance was found with older participants having a lower error score (r_s = -.26, p < .0005). When dividing the sample by gender, the correlation was significant for male participants (r_s = -.38, p < .0005), but not for female participants (r_s = -.17, p = .09; z_{r1-r2} = 1.54, p = .12).

No association was observed between IQ and Sentence Completion Task performance with r_s values ranging from -.03 to -.08 between the error score and all IQ indices (all p > .26). Furthermore, no correlations were significant within male (all r_s < -.11, p > .28) or female participants (all r_s < -.09, p > .39).

8.2.5 ASD and Control Results

Table 8.2 presents group means for the three measures of performance on the Sentence Completion Task for the ASD and control groups (excluding one control participant with ESL). Participants in both groups took significantly longer to respond to test items than control items (Wilcoxon's Signed Ranks Tests, z > 3.18, p < .002).

Furthermore, ASD and control groups did not differ in the mean time to complete each sentence stem for test items ($z = 1.11, p = .27$) or control items ($z = 0.97, p = .33$).

One-tailed Mann-Whitney *U* tests found the ASD group made significantly more local completions ($z = 2.63, p = .004$) and had a higher error score ($z = 2.88, p = .002$) than the control group. This difference held for the high-functioning subgroups (local completions $z = 2.29$, $p = .01$; error score $z = 2.76$, $p = .003$). A trend towards a significant group difference was found between the low-functioning subgroups on the number of local completions made ($z = 1.48, p = .07$) and error score ($z = 1.30, p = .10$).

In a frequency analysis it was shown that 45% of the ASD group (14 of 31 participants) had an error score of at least 4, while this was true for only 20% of the control group (6 of 30 participants; $\chi^2 = 4.38$, $p = .04$). Furthermore, most participants in the control group (67%) never made a local completion, but the majority of participants in the ASD group (65%) made one or more such local answers ($\chi^2 = 5.93, p = .01$).

Table 8.2 Sentence Completion Task results for each group overall, and split by high and low IQ: Mean (*SD*).

	Group	Error score[a]	Number of local completions[b]	Response time per completion (seconds)
All	ASD	3.97 (2.88)	1.35 (1.40)	3.77 (2.23)
	Control	2.19 (2.63)	0.61 (1.17)	3.38 (2.53)
High IQ	ASD	3.56 (2.45)	1.12 (1.20)	3.91 (2.35)
	Control	1.92 (2.55)	0.52 (1.08)	3.16 (2.55)
Low IQ	ASD	5.67 (4.08)	2.33 (1.86)	3.18 (1.66)
	Control	3.33 (2.88)	1.00 (1.55)	4.32 (2.39)

[a]ASD > control, $p = .002$; $\text{ASD}_{\text{high}} > \text{Control}_{\text{high}}$, $p = .003$
[b]ASD > control, $p = .004$; $\text{ASD}_{\text{high}} > \text{Control}_{\text{high}}$, $p = .01$

Correlations

The pattern of correlations did not suggest a speed-accuracy trade-off in completing sentences in a local manner. In the ASD group, a non-significant negative correlation ($r_s = -.29, p = .12$) was found between the number of local responses and time to complete each sentence, while no association was found between the two variables in the control group ($r_s = .07, p = .71$; $z_{r1-r2} = 1.34, p = .18$).

The correlation between age and error score did not reach significance in the control group ($r_s = -.26$, $p = .16$), but was of similar magnitude as found in the TD sample. No such association was found in the ASD group ($r_s = .13$, $p = .47$), although the coefficients did not differ significantly in magnitude between groups ($z_{r1\text{-}r2} = 1.56$, $p = .12$). Participants with lower MA tended to make more local errors on the Sentence Completion Task in both groups, although the correlation coefficients did not reach statistical significance (ASD $r_s = -.31$, $p = .10$; control $r_s = -.30$, $p = .10$).

A significant negative correlation was observed between FIQ and the error score in the ASD group ($r_s = -.38$, $p = .03$), which was not found in the control group ($r_s = -.12$, $p = .52$; $z_{r1\text{-}r2} = 1.04$, $p = .30$). Similarly, VIQ and FIQ-BD showed some association with the error score in the ASD group ($r_s = -.34$ and $-.35$ respectively, $p > .05$) which, again, was not observed in the control group ($r_s = .02$ and $-.05$, $p > .79$; $z_{r1\text{-}r2} < 1.39$, $p > .16$). Participants with high PIQ tended to have lower error scores in both groups, although neither correlation reached statistical significance (ASD $r_s = -.27$, $p = .15$; control $r_s = -.25$, $p = .18$).

8.2.6 Summary of Sentence Completion Task

Local completions to sentence stems, along with prolonged hesitations, were more common in younger children in the TD sample. Gender effects were found, with a trend for males to provide local completions more often than females. The effect of age was also stronger in male participants. Furthermore, a trade-off between speed and accuracy was evident in female participants only.

As predicted from weak central coherence theory, individuals with ASD were more likely to provide a local completion to sentence stems, rather than a globally coherent response, compared to their age and IQ-matched control group. A trend was found for participants with low MA in both groups to give local completions and/or take longer to provide a response. Lower intelligence was associated with more local completions in the ASD group, while these variables appeared to be independent in both the control group and the TD sample.

The error score taken from the Sentence Completion Task was used as a continuous measure of central coherence, with higher scores indicating weaker coherence. Participants were categorised as showing weak central coherence if they gave two or more local completions on the task.

8.3 Homograph Reading Task

A major obstacle to the comprehension of linguistic material is that single words can often have more than one meaning. To resolve any ambiguity, words are interpreted according to their presented context, a process that is relatively automatic in TD individuals. This "drive for meaning" implies that many ambiguous words such as homophones, words that have the same sound but different meanings (e.g., pear, pair), are hardly noticed in everyday speech. Similarly, the interpretation of homographs, words that have the same spelling but differ in meaning and often pronunciation (e.g., *bow* the front of a ship, and *bow* a loop made in string), are disambiguated by a given context.

Homograph Reading in ASD

According to the central coherence hypothesis, individuals with autism would tend not to read connected text for meaning, perhaps failing to integrate words in a sentence. Frith and Snowling (1983) explored this hypothesis in a series of experiments. When asked to read a short story aloud that included homographs, they found children with autism tended to give the more frequent pronunciation of the homograph irrespective of sentence context. In contrast children with dyslexia, matched in reading ability to the children with autism, would rely on context in order to decode unknown words.

Subsequent studies have consistently found a deficit in autism in the use of contextual information within sentences to disambiguate homographs (Happé, 1997; Jolliffe & Baron-Cohen, 1999; Lopez & Leekam, 2003). Happé (1997) adapted materials originally used by Snowling and Frith (1986), to study the influence of position of the homograph in the sentence. All children were more accurate when the homograph followed the context, than when it appeared initially, but TD children were more sensitive to this effect of position than children with autism. Jolliffe and Baron-Cohen (1999) replicated the finding that people with autism make reduced use of preceding sentence context to guide the correct pronunciation of homographs, this time in a sample of high-functioning adults with ASD. Although the effect of homograph position was less apparent in their sample, they reasoned that older readers might be more able to read ahead.

Homograph Reading in Typical Development

The reading of homographs has been well studied in the general literature, primarily as a means to investigate age-related changes in the use of context. An established finding is that young children depend heavily on context to decipher words, and context-effects of priming on word identification decrease with age (Simpson & Lorsbach, 1983). Word frequency effects are said to be more established in older children, with the more common interpretation of a homograph having greater activation, over and above the effects of context (Simpson, 1981). However the role of context versus word-frequency in resolving ambiguous words is much debated (Dixon & Twilley, 1999; Gernsbacher & Faust, 1991).

Developmental effects on homograph reading have been studied extensively by Simpson and colleagues (Simpson & Foster, 1986; Simpson & Kreuger, 1991; Simpson, Krueger, Kang, & Elofson, 1994). In these studies participants are presented with sentences that end in a homograph, with the context biased towards either the frequent or rare meaning of the homograph. A target word follows that is either an associate to the frequent or rare meaning of the homograph, and time to read the target word aloud is taken as the dependent variable. Using this paradigm, 9-year-old children were greatly influenced by the biasing context, more so than 12-year-old children. The 12-year-olds were, however, more sensitive to the frequency of the homograph, with greater facilitation effects when the preceding context was biased towards the frequent meaning, more so than the rare meaning. In contrast, response times for 9-year-olds were similar whether the word was biased toward the frequent or rare meaning. This lack of frequency effect was also found in adults (Simpson & Kreuger, 1991), and it was reasoned that qualitative differences in reading exposure at different ages could account for this finding. Due to lower reading experience, young children may not have established a hierarchy of word frequencies. Adults on the other hand, have greater experience with both meanings and can be more flexible between the two interpretations, showing less reliance on word frequency. The 12-year-olds are considered to have enough experience with words to be influenced by the dominant meaning only.

This developmental effect on homograph frequency has also been demonstrated in studies that have examined the activation of alternative meanings of homographs when presented in isolation. Simpson and Foster (1986) presented 8-, 10- and 12-year-old children with homographs (e.g., bark), followed by either a frequent associate (e.g., dog) or rare associate (e.g., tree) which participants were required to read aloud. When the subsequent word was presented very quickly (150 ms or 300 ms), both homographs meanings were equally activated. It was not until target words were presented 750 ms after the homograph that the dominant meaning was activated first, and this was only found in 12-year-olds. Word frequency effects may therefore only be seen in the present study in TD participants older than 12 years of age.

Predictions

The Homograph Reading Task as described in Happé (1997) was used in the present study as a measure of coherence in the verbal-semantic domain. Individual differences were measured through the effect of context position; participants who use sentence context to derive the appropriate pronunciation of homographs should be sensitive to the position of the target word relative to the sentence context. Context effects are expected to be most clearly demonstrated when the rare pronunciation of the homograph is anticipated. Individuals with ASD were predicted to show a failure to use context and provide the frequent pronunciation of the homograph regardless of context. They are also expected to show a lower tendency to correct initial errors compared to control participants.

Although the processes relevant to reading homographs in context (e.g., use of context to decode words, awareness of word-frequency) are suggested to change with development, no specific predictions are made regarding the developmental course of context-position effects in TD because these processes may develop independently and a slightly different ages.

8.3.2 Method

Materials

The stimuli consisted of 16 test sentences and a set of 13 pre-test single words that included the homographs. Each test sentence appeared on a separate A4 page, printed across a single line (landscape orientation), while the pre-test words appeared in a column on a single page (portrait orientation).

There were four types of sentences originating from Snowling and Frith (1986) and further adapted by Happé (1997): two sentences in which the more frequent pronunciation of the homograph was correct and two sentences in which the rare pronunciation was appropriate. For each pronunciation the homograph was positioned once near the beginning of the sentence, before the disambiguating context, and once at the end of the sentence, after the disambiguating context.

Four of the original five homographs from Happé (1997) were used (*tear*, *row*, *lead*, *bow*). The homograph *read* was excluded as the grammatical structure of the sentence can be used to disambiguate the homograph rather than the semantic context. It has been shown that children with autism can correctly use syntactic information in sentences (Frith & Snowling, 1983).

For the adult participants assessed in New Zealand the homograph *tear* was replaced by another homograph (*invalid*) as the two pronunciations of the phoneme /*ea*/ are not always differentiated in the New Zealand accent. The frequent pronunciation (i.e., not valid) and rare pronunciation (i.e., sick person) of *invalid* were confirmed from the norms collected by Twilley, Dixon, Taylor, and Clark (1994). The new sentences were: The old invalid passport barred him from entry (before context, frequent pronunciation); The old invalid dragged himself to bed (before context, rare pronunciation); His expired driver's licence was now invalid (after context, frequent pronunciation); After suffering injuries in the war he was placed on the invalid pension (after context, rare pronunciation).

To assess participants' knowledge of each homograph, for both meaning and pronunciation, a post-test in the form of a picture-vocabulary test was developed. The post-test consisted of a 10-page booklet. On each page four black and white line drawings were presented, one of which depicted the meaning of a homograph printed at the bottom of the page. Each definition of the homographs portrayed in the test sentences was assessed: *bow:* weapon with arrows, musical device, bend body; *lead:* to guide, strap for animal, metal; *row:* line, argue; *tear:* liquid from eye, rip.

Procedure

Each participant was asked to read aloud the pre-test word list followed by the test sentences. The pre-test word list was presented first to check the participant's reading ability and familiarity with the target words. Four participants (one ASD, three low-functioning control participants) had difficulty reading the words and were not administered the test sentences.

Test sentences were presented to the participant one at a time in a fixed pseudo-random order with no two homographs appearing consecutively. The experimenter corrected any reading errors made at the time of reading, other than on the homograph targets, in order to assist comprehension of the text. The experimenter marked on a record sheet the initial pronunciation of each homograph plus any self-corrections made by the participant.

Post-test

As Lopez and Leekam (2003) noted, a limitation of the homograph task as used by Happé (1997) and Jolliffe and Baron-Cohen (1999), is that baseline knowledge of the rare pronunciations of the homographs is not assessed and the assumption is made that participants are familiar with both uses of the target words. As Happé (1997) comments, it is necessary to avoid directing participants to the special status of the homographs before or during the reading task, in order to make the task as naturalistic as possible. A measure of homograph familiarity was therefore administered at a later date to the homograph test (either at the end of the second or third session), to avoid alerting participants to the double meanings and pronunciations of the homographs at the time of testing. All children were administered the post-test, while participants in the older group (17-25 years) were given the post-test if they had English as a second language, or if test performance led the experimenter to doubt the participant's knowledge of a homograph.

Participants were presented with the post-test booklet and were instructed to first look at the target word, select the picture that best showed the meaning of the word, and then read the word aloud. If the participant selected an incorrect picture and incorrectly pronounced the word, the experimenter provided the correct pronunciation and asked which picture best showed the meaning of this new word. If an incorrect picture was selected but was paired with a correct pronunciation, the experimenter asked why that picture was chosen. Similarly, if the correct picture was chosen, but an incorrect pronunciation was provided, the experimenter asked why the particular selection was made.

This follow-up was necessary to confirm whether the participant did have semantic knowledge of the homograph but an incorrect association with pronunciation.

8.3.3 Scoring

Participants had to demonstrate that they knew all meanings and pronunciations for at least two of the four homographs on the post-test for their results (on the relevant test sentences) to be considered valid. For participants who showed knowledge of only two or three homographs in the post-test, total scores were prorated from valid item sentences to aid comparisons.

Participants were scored on their correct pronunciations on the test sentences, which included both initial and corrected attempts. Each participant received a total score out of 16 across the whole task and a score out of four for each of the four sentence conditions. On occasion, participants changed an initially correct pronunciation to an incorrect one. This happened on ten occasions in the TD sample (once for five participants in the 8-10 year group, once for three participants in the 11-13 year group, and twice for one participant in the 14-16 year group) and on two occasions in the ASD group (once for two participants). The participant's final response was scored and their familiarity with the homograph was checked in the post-test.

8.3.4 Typical Development Results

Four participants (one male from the 14-16 year group, two female and one male from 17-25 year group), all of whom had English as a second language, failed the post-test criterion and were excluded from analysis. The removal of these participants did not impact on the pattern of IQ performance between age groups. The mean accuracy scores for the four age groups, by sentence type, are presented in Table 8.3 for initial and final responses. The data were extremely negatively skewed in each age group (all z-scores of skewness > -2.8), implying that participants were reading the majority of homographs true to context. Due to the non-normality of the data non-parametric statistical tests were used.

Table 8.3 Number of homographs pronounced context-appropriately for initial and final responses (maximum = 4) by age group: Mean (*SD*).

Age group (years)	N		Frequent pronunciation		Rare pronunciation	
			Before[a]	After[b]	Before	After[c]
8-10	54	initial	2.07 (1.18)	3.36 (0.86)	2.43 (1.27)	2.99 (1.03)
		final	*3.16 (1.02)*	*3.56 (0.85)*	*3.34 (1.08)*	*3.34 (0.98)*
11-13	43	initial	2.33 (1.12)	3.36 (0.81)	2.46 (1.09)	3.34 (0.88)
		final	*3.58 (0.80)*	*3.61 (0.71)*	*3.49 (0.86)*	*3.61 (0.79)*
14-16	50	initial	3.03 (0.99)	3.76 (0.43)	2.77 (1.13)	3.57 (0.76)
		final	*3.73 (0.58)*	*3.90 (0.30)*	*3.69 (0.74)*	*3.77 (0.72)*
17-25	53	initial	3.02 (0.96)	3.59 (0.67)	2.77 (1.06)	3.45 (0.74)
		final	*3.73 (0.54)*	*3.81 (0.48)*	*3.68 (0.73)*	*3.56 (0.70)*

[a] initial 8-10, 11-13 < 14-16, 17-25, $p < .004$; final 8-10 < 11-13, 14-16, 17-25, $p < .03$
[b] initial 8-10, 11-13 < 14-16, $p < .02$
[c] initial 8-10 < 14-16, 17-25, $p < .01$; final 8-10, 17-25 < 14-16, $p < .03$

A developmental progression was shown in the number of homographs correctly pronounced (allowing for self-corrections) across all 16 sentences ($\chi^2 = 22.51$, $p < .0005$), with a significant increase from 8-10 ($M = 13.40$, $SD = 2.28$) to 11-13 years ($M = 14.30$, $SD = 2.23$; $z = 2.15$, $p = .03$); and from 11-13 to 14-16 years ($M = 15.10$, $SD = 1.69$; $z = 2.17$, $p = .03$). No difference was found between the two oldest age groups (17-25 years, $M = 14.78$, $SD = 1.62$; $z = 1.58$, $p = .11$). Male and female participants performed equivalently in the total number of homographs read correctly ($z = .02$, $p = .98$).

Effects of age group were found in the four sentence conditions of the Homograph Reading Task. When taking the initial pronunciation made by participants, and not allowing for self-corrections, a significant effect of age group was found for frequent homographs, presented either before ($\chi^2 = 28.06$, $p < .0005$) or after context ($\chi^2 = 9.29$, $p = .03$; see Table 8.3). No age group effects were detected for the initial pronunciation of rare homographs when presented before the context ($\chi^2 = 3.95$, $p = .27$), but were found when the homographs were presented after the context ($\chi^2 = 12.83$, $p = .005$).

When taking participants' final response and allowing for self-corrections, a significant effect of age group was found when the more frequent pronunciation of the homograph was expected and was placed before the context ($\chi^2 = 16.24$, $p = .001$), and approached significance when placed after the context ($\chi^2 = 7.53$, $p = .06$). When the rare pronunciation of the homograph was expected, age groups did not differ when the homograph was presented before the context ($\chi^2 = 5.76$, $p = .12$), but did show a significant effect when the homograph followed the context ($\chi^2 = 9.93$, $p = .02$; see Table 8.3).

There was no main effect of gender on the number of correctly pronounced homographs (initial and final responses) in any of the four sentence conditions (all $z < 1.61$, $p > .10$).

The 17-25 year group did not outperform the 14-16 year group in any condition, and performed significantly lower than the 14-16 year group for the final pronunciation of rare homographs presented after context ($z = 2.28$, $p = .02$). This effect may have arisen from the high number of participants for whom English was a second language in the oldest age group (N = 5 who had not been excluded as a result of the post-test). The above analyses were repeated after excluding participants from all age groups who had ESL. As this exclusion did not affect the pattern of significant results, these participants remained in the analysis.

As the majority of participants in the 17-25 year group were from the New Zealand sample (N = 30) it was considered that the use of the homograph *invalid* might have related to the relatively inferior performance of this age group. Indeed, only 63% of the sample was familiar with the rare pronunciation of invalid. The above analyses were repeated after excluding this homograph in the oldest age group and prorating each participant's score from the remaining three homographs. Excluding this homograph did not alter the pattern of significant results and so all data were retained in further analyses.

Position of homograph

Central coherence was measured by the effect of context position; individuals with strong coherence are expected to show greater sensitivity to the position of the target homograph relative to the sentence context. Wilcoxon's Signed Ranks Tests were therefore used to assess the effect of homograph position for individuals within each age group. For initial pronunciations of both rare and frequent homographs, all age groups gave significantly more correct pronunciations when the context preceded versus followed the homograph (all $z > 2.63$, $p < .009$). For final responses, this effect was apparent in the 8-10 and 14-16 year groups when the frequent pronunciation of the homograph was appropriate ($z = 2.69$, $p = .008$; $z = 2.00$, $p = .05$, respectively).

When analysing by gender, position effects were found for both male and female participants for the initial pronunciations of rare and frequent homographs (all $z > 4.79$, $p < .0005$). Significant effects of position for the final pronunciation of frequent homographs were only found in male participants ($z = 3.09$, $p = .002$), and not female ($z = 1.49$, $p = .14$). When analysing by age group, this effect of position was found for males in the oldest age group ($z = 2.12$, $p = .03$), and approached significance for males in the 14-16 year group ($z = 1.85$, $p = .06$).

A difference score was calculated to measure individual effects of sentence context for initial pronunciations, across all homograph types (i.e., the number of correctly pronounced homographs presented the before context subtracted from the number correct when presented after context). The difference score was calculated for both rare and frequent homographs as it could not be determined whether all children were sensitive to word frequency effects. No overall effect of age group was found on the difference score ($\chi^2 = 4.89$, $p = .18$), although males had a higher difference score ($M = 1.90$, $SD = 1.65$) compared to females ($M = 1.39$, $SD = 1.55$; $z = 2.02$, $p = .04$). This gender effect was found in the 14-16 year group ($z = 2.16$, $p = .03$) and approached significance in the 11-13 year group ($z = 1.86$, $p = .06$). This effect of position is shown in Figure 8.2 and Figure 8.3 by the steeper slope for male participants in these age groups.

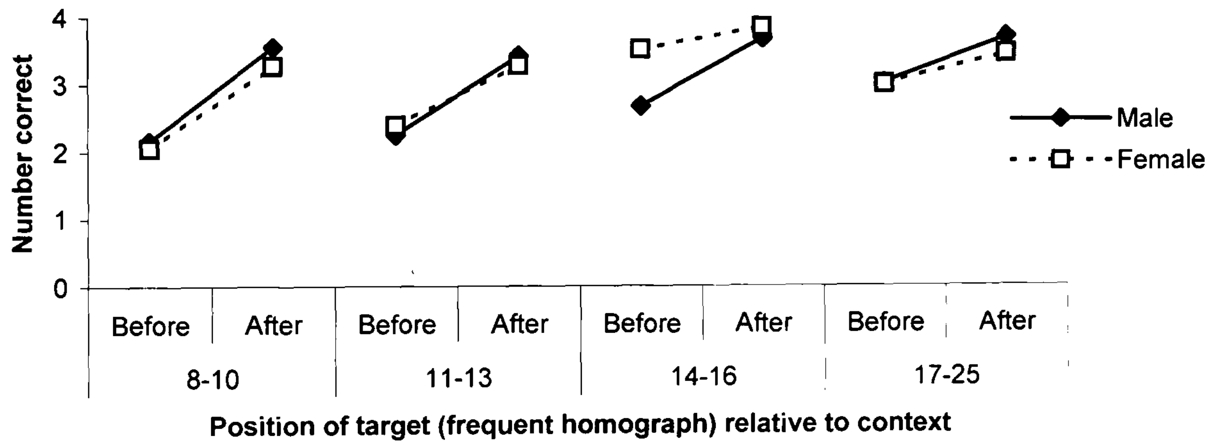


Figure 8.2 Effect of target position on initial pronunciation of frequent target words for male and female participants in each age group.

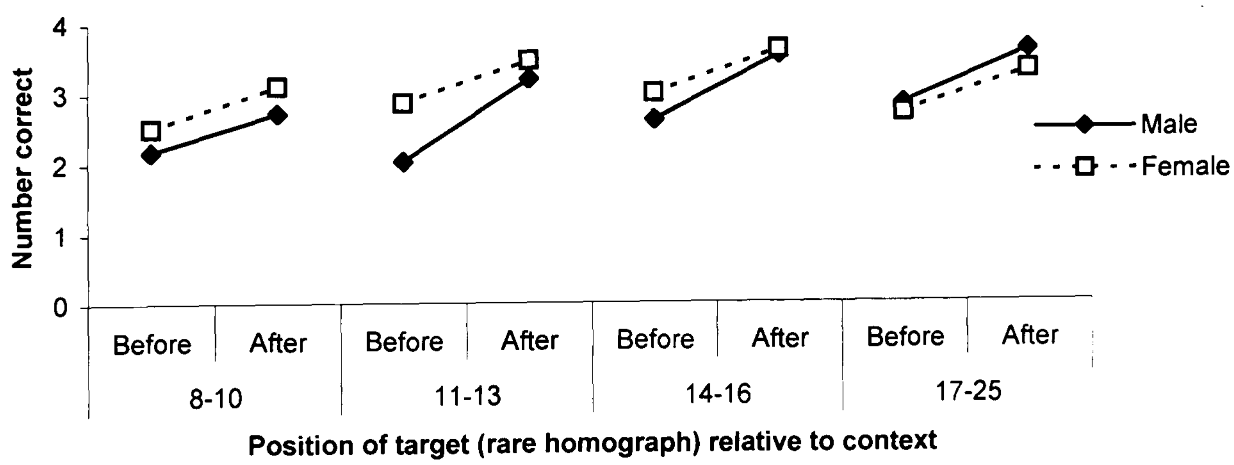


Figure 8.3 Effect of target position on initial pronunciation of rare target words for male and female participants in each age group.

Self-corrections

Participants who made at least one error in their initial pronunciation of a homograph were compared on their proportion of self-corrections (number of self-corrections/number of initial errors). A significant effect of age group was found on this variable ($\chi^2 = 12.3$, $p = .006$), with the 8-10 year group making significantly fewer self-corrections ($M = 0.58$, $SD = 0.36$) than the 11-13 year ($M = 0.71$, $SD = 0.33$) and 14-16 year groups ($M = 0.81$, $SD = 0.31$; $z > 1.92$, $p < .05$). The 14-16 year group made significantly more self-corrections than the oldest age group ($M = 0.67$, $SD = 0.38$; $z = 1.93$, $p = .05$). Male and female participants made an equivalent proportion of self-corrections overall ($z = 0.19$, $p = .85$), but as shown in Figure 8.4, females self-corrected significantly more than males in the 14-16 year group ($z = 2.23$, $p = .03$). As female participants had significantly higher VIQ than males in this age group, this verbal advantage may have contributed to this gender effect.

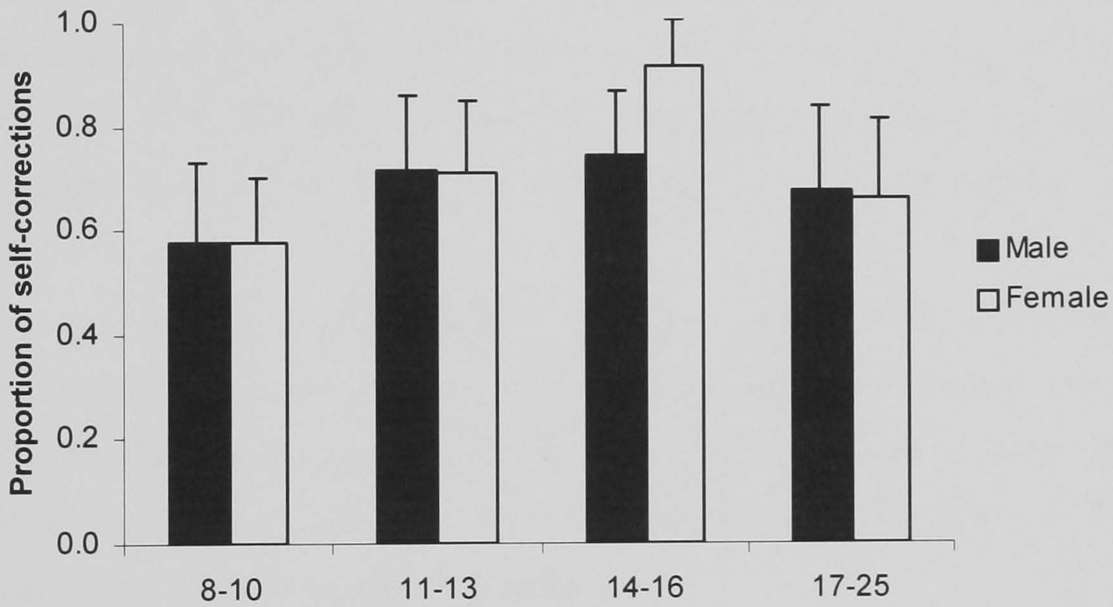


Figure 8.4 Proportion of self-corrections made by each age group split by gender. Error bars show upper 95% confidence intervals.

Correlations

The difference in number of homographs read correctly on the initial pronunciation, when the homograph was positioned after the sentence context compared to before, was used as the main coherence index for correlational analyses. This difference score showed a significant negative correlation with age in the entire TD sample ($r_s = -.16$, $p = .03$) suggesting that older participants showed less effect of context. The correlation was significant in female participants ($r_s = -.19$, $p = .05$), but did not reach significance for male participants ($r = -.16$, $p = .11$; $z_{r1-r2} = 0.19$, $p = .85$).

There was no association between IQ and the difference score across the whole sample ($r_s = -.06$ to $-.08$, $p > .23$). This was also evident for female participants ($r_s = -.02$ to

.03, $p > .76$). In contrast, male participants showed a negative correlation between the difference score and IQ, which was significant for FIQ and PIQ ($r_s = -.21$, $p = .04$), and approached significance for VIQ and FIQ-BD ($r_s = -.19$, $p = .07$). Fisher's z transformation did not find the strength of correlation coefficients to differ significantly between genders (all $z_{r1-r2} < 1.60$, $p > .10$).

A significant positive correlation was found between the proportion of self-corrections and age in the TD sample ($r_s = .15$, $p = .04$), but did not reach statistical significance when analysing by gender (males $r_s = .14$, $p = .18$; females $r_s = .16$, $p = .11$). The proportion of self-corrections correlated positively with all IQ measures ($r_s = .21$ to .34, males $r_s = .24$ to .37, females $r_s = .20$ to .32, all $p < .05$).

8.3.5 ASD and Control Results

Table 8.4 presents the results from the Homograph Reading Task for the ASD and control groups. Data from one ASD participant and three control participants were removed from analysis because of failure on the post-test criterion. All four participants were classed as having low IQ and their exclusion did not alter the group matching on age and IQ (all $t_{(56)} < 0.70$, $p > .48$). As the data were negatively skewed for both groups (z-scores ranged from -0.14 to -4.13), nonparametric statistical analyses were used throughout.

Overall performance was equivalent between groups, with ASD and control participants pronouncing the same number of homographs correctly overall (ASD $M = 13.36$, $SD = 2.73$; control $M = 14.38$, $SD = 2.00$; $z = 1.48$, $p = .14$, two-tailed). Comparing groups with high IQ ($z = 1.20$, $p = .23$) and low IQ ($z = 1.36$, $p = .17$) also did not reveal any differences in overall performance.

Group comparisons were made on each of the four sentence conditions. In the analysis of initial pronunciations, the ASD group pronounced more frequent homographs correctly when presented before the context than the control group ($z = 2.56$, $p = .01$, two-tailed). When allowing for self-corrections, the ASD group showed a trend in giving fewer context-appropriate rare pronunciations than the control group when the homograph was positioned before the context ($z = 1.78$, $p = .07$, two-tailed). This pattern of group performance was observed between groups of high IQ participants (see Table 8.4), but not between groups of low IQ participants (although the number of participants who passed the post-test criteria was low in the latter group).

Table 8.4 Number of homographs pronounced context-appropriately (max = 4) by group: Mean (*SD*).

				Frequent pronunciation		Rare pronunciation	
	Group	N		Before[a]	After	Before[b]	After
All	ASD	30	initial	3.09 (0.93)	3.58 (0.59)	2.08 (1.41)	3.08 (0.97)
			final	*3.52 (0.64)*	*3.72 (0.47)*	*2.76 (1.39)*	*3.36 (1.01)*
	Control	28	initial	2.43 (1.01)	3.42 (0.72)	2.56 (1.18)	3.30 (0.92)
			final	*3.50 (0.77)*	*3.76 (0.47)*	*3.45 (0.80)*	*3.67 (0.68)*
High IQ	ASD	25	initial	3.28 (0.74)	3.67 (0.50)	2.16 (1.40)	3.23 (0.89)
			final	*3.72 (0.46)*	*3.80 (0.41)*	*2.93 (1.36)*	*3.52 (0.92)*
	Control	25	initial	2.48 (1.02)	3.39 (0.74)	2.71 (1.15)	3.41 (0.79)
			final	*3.56 (0.74)*	*3.73 (0.49)*	*3.63 (0.59)*	*3.79 (0.53)*
Low IQ	ASD	5	initial	2.13 (1.26)	3.13 (0.87)	1.67 (1.49)	2.33 (1.11)
			final	*2.53 (0.51)*	*3.33 (0.62)*	*1.87 (1.26)*	*2.53 (1.12)*
	Control	3	initial	2.00 (1.00)	3.67 (0.58)	1.33 (0.58)	2.33 (1.53)
			final	*3.00 (1.00)*	*4.00 (0.00)*	*2.00 (1.00)*	*2.67 (1.15)*

[a]initial ASD > Control, p = .01 $ASD_{high} > Control_{high}$, p = .005
[b]final ASD < Control, p = .07 $ASD_{high} < Control_{high}$, p = .10

Position of homograph

As group differences were predicted in the effect of homograph position, Wilcoxon's Signed Ranks Tests were performed to assess individual effects of homograph position within each group. Position effects were found in all sentence conditions with significantly more homographs read correctly when they followed, rather than preceded, the context. This effect was observed for the initial pronunciations of frequent homographs (ASD z = 2.16, p = .03; control z = 3.30, p = .001; see Figure 8.5) and rare homographs (ASD z = 3.59, p < .0005; control z = 2.92, p = .004; see Figure 8.6); and for the final pronunciations of rare homographs (ASD z = 3.11, p = .002; control z = 2.12, p = .03). The effect of position for the final pronunciation of frequent homographs approached significance for the ASD group (z = 1.73, p = .08), but was significant in the control group (z = 1.98, p = .05).

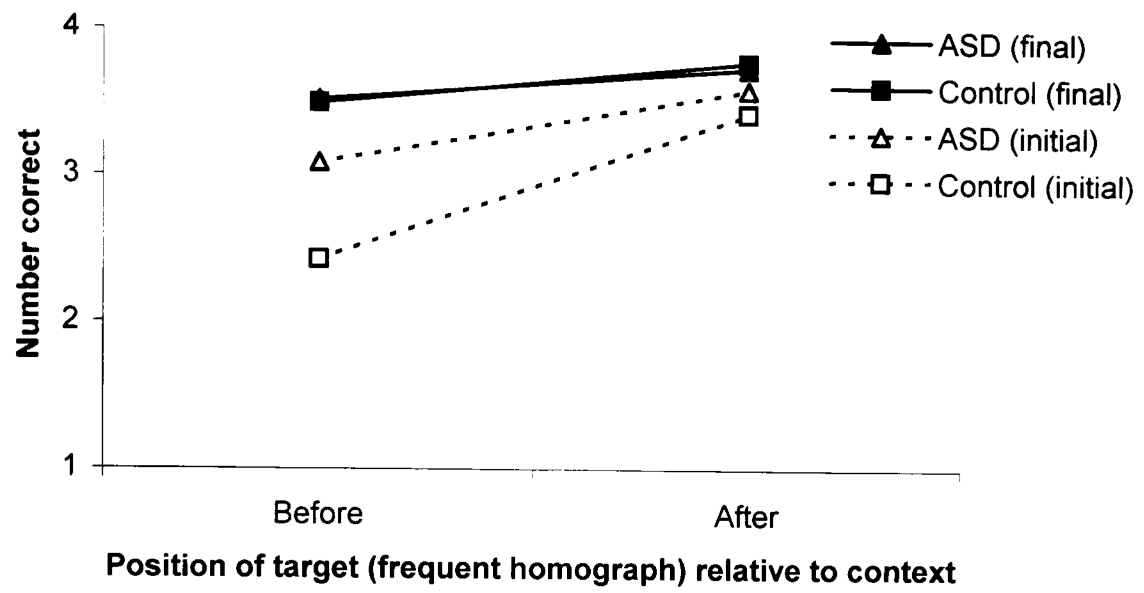


Figure 8.5 Effect of target position on correct pronunciation of frequent target words.

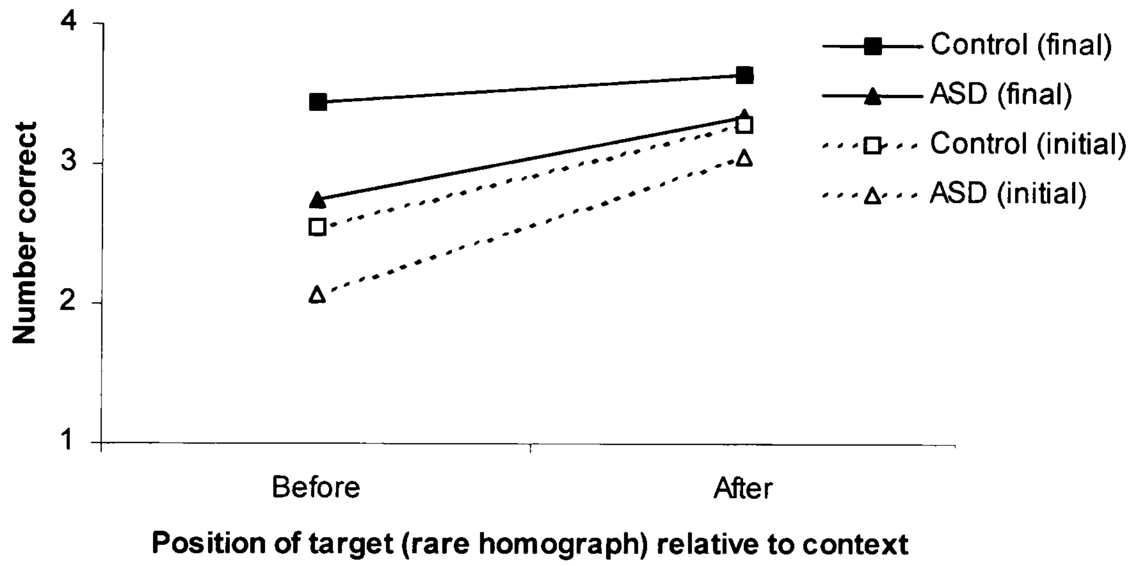


Figure 8.6 Effect of target position on correct pronunciation of rare target words.

Individual effects of sentence context were assessed by comparing the difference in the number of correct initial pronunciations for homographs presented after compared to before the sentence context. This difference score was higher in the control group ($M = 1.73$, $SD = 1.76$) than the ASD group ($M = 1.49$, $SD = 1.45$), suggesting that sensitivity to sentence context was more apparent in the control group, but the difference was not significant ($z = 0.85$, $p = .42$, one-tailed). No differences were found between groups when only considering participants with high IQ ($z = 0.74$, $p = .23$), or low IQ ($z = 0.60$, $p = .27$).

Self-corrections

Groups were equivalent on the mean number of initial errors made by each participant over the entire task (ASD $M = 4.00$, $SD = 2.33$; control $M = 4.14$, $SD = 2.46$; $z = 0.26$, $p = .80$), however the ASD group were less likely to spontaneously correct an error. For those participants who ever made an error (ASD $n = 29$; control $n = 27$), the proportion of self-corrections was significantly lower for the ASD group ($M = 0.47$, $SD = 0.42$) than the control group ($M = 0.71$, $SD = 0.31$; $z = 2.22$, $p = .01$, one-tailed). This difference was significant when including only participants with high IQ (ASD $M = 0.53$, $SD = 0.43$, $n = 24$; control $M = 0.76$, $SD = 0.30$, $n = 24$; $z = 1.84$, $p = .03$), and approached significance for participants with low IQ (ASD $M = 0.20$, $SD = 0.19$, $n = 5$; control $M = 0.35$, $SD = 0.06$, $n = 3$; $z = 1.34$, $p = .09$).

In a frequency analysis, 96% of the control group (26 out of 27), if they ever made an error, self-corrected at least once; while only 68% of the ASD group (20 out of 29) who performed below ceiling and had the opportunity to correct themselves, ever did so. The ASD group corrected fewer of their errors than did the 8-10 year group from the TD sample (see Figure 8.4).

Correlations

The difference score calculated to measure sensitivity to sentence context for initial pronunciations was not associated with age in the ASD group ($r_s = .02$, $p = .94$) or control group ($r_s = -.12$, $p = .56$). The difference score showed a negative correlation with MA in the control group, which approached significance ($r_s = -.36$, $p = .06$). No such association was found in the ASD group ($r_s = -.03$, $p = .88$; $z_{r1-r2} = 1.26$, $p = .20$).

A significant negative correlation was found in the control group between the difference score and IQ (FIQ $r_s = -.44$, $p = .02$; FIQ-BD $r_s = -.44$ $p = .02$; VIQ $r_s = -.46$, $p = .01$; with the exception of PIQ $r_s = -.31$, $p = .11$). In contrast, the difference score did not correlate with any IQ measure in the ASD group ($r_s = -.12$ to $-.004$, all $p > .53$). The difference in magnitude of the correlation coefficients between the ASD and control groups did not reach significance (all $z_{r1-r2} < 1.79$, $p > .08$).

The proportion of self-corrections did not correlate with age in the ASD ($r_s = .14$, $p = .46$) or control group ($r_s = -.17$, $p = .40$), but was positively associated with MA in both groups (ASD $r_s = .40$, $p = .03$; control $r_s = .37$, $p = .06$). The proportion of self-corrections correlated significantly with all IQ measures in the control group ($r_s = .54$ to $.66$, all $p < .003$), but not significantly in the ASD group ($r_s = .22$ to $.36$, all $p > .05$; $z_{r1-r2} < 1.70$, $p > .10$).

8.3.6 Summary and brief discussion of Homograph Reading Task

In summary, overall reading accuracy for homographs presented in context was found to increase with development in the TD sample up until 14-16 years. Context effects in the initial reading of homographs were stronger in younger TD participants. A gender effect was found with male participants showing a stronger effect of context than female participants. IQ effects were detected only in male participants, with those of lower IQ showing a greater effect of context position.

The ASD group was expected to show a reduced use of sentence context to inform the correct pronunciation of homographs. The ASD and control groups were equivalent in homograph reading performance, in overall correctness and the number of initial errors made. However, the ASD group were less likely to correct an initial error compared to the control group, especially when the rare pronunciation of the homograph was expected. The main indicator of weak coherence was predicted to be reduced effect of context position; however no group differences were detected on this measure. As found for TD males, sensitivity to context position was influenced by IQ in the control group, but this association was not apparent in the ASD group. The proportion of self-corrections was also driven by IQ in the control group, more so than in the ASD group.

A reduced use of sentence context in the ASD group was only weakly supported, evidenced by fewer self-corrections made compared to the control group. Several considerations may explain why group differences were not found for the effect of context position. The observation that older TD participants, and males with high IQ (TD and control) tended to show less effect of context could have resulted from ceiling effects, possibly reflecting reading ahead. In support of this, Jolliffe and Baron-Cohen (1999) also found a lack of an effect of position of the homograph in their TD adults (18 to 49 years). They proposed that older experienced readers might be able to read ahead more easily than younger normal readers, and hence show no difference whether the homograph precedes or follows the sentence context.

A second issue to consider in the present study is the age-related effects on word-frequency familiarity. It was found that young TD children did not automatically say the more frequent pronunciation of the homograph when the preceding context did not guide either pronunciation, which suggests that frequency effects may only be apparent when an individual is exposed to written language for a number of years (i.e., beyond 13 years of age). As detailed in the TD literature, age-related changes may occur in the degree to which context is used, but also in the hierarchical influence of word frequency (Simpson & Kreuger, 1991; Simpson et al., 1994). Normative studies have reported that the age of acquisition for various homographs can also differ (Gilhooly & Logie, 1980), although more up-to-date norms for UK based populations are needed. Information regarding the familiarity of each homograph and associated frequency in relation to age would be essential in the design of future studies.

A limitation of the current study was found in the later analysis of the presentation order of sentences. All participants received the sentences in a pseudo-random order.

Inadvertently, however, more sentences containing rare homographs were administered in the first half of the task (6/8), while more sentences containing frequent homographs were administered in the final half (6/8). It is plausible that prior exposure to the rare pronunciation influenced the subsequent readings of frequent homographs, reducing the expected word frequency and context position effects.

Two candidate variables were taken forward as indices of coherence: (1) the effect of target position (after minus before) on the initial pronunciation of target homographs and (2), the proportion of self-corrections. Higher scores on both indices indicated a greater awareness of sentence context, and therefore stronger central coherence in the verbal domain. Participants were categorised as showing weak central coherence if two or more errors were made when reading rare homographs, which were not corrected.

8.4 Story Memory Task

Soon after a sentence has been comprehended, information about its exact surface form (e.g., word order) readily fades (Bransford & Franks, 1971; Gernsbacher, 1985; Sachs, 1967). As a consequence, memory for linguistic information is rarely word for word, although the general meaning is typically preserved. Individuals with autism have not always shown this phenomenon, demonstrating strengths in rote memory and in the exact recall of sentences (Aurnhammer-Frith, 1969; Hermelin & O'Connor, 1967). As good rote memory is often accompanied by poor ability to extract meaning, Frith (2003) proposed that the tendency to recall information verbatim might be a valid indicator of weak central coherence. In order to test this hypothesis, a measure of story memory was included in the present study. The tendency for gist versus verbatim recall was assessed, in addition to recognition memory for the surface form of sentences.

Verbal memory in Typical Development

Age-related improvements in verbal memory are observed in TD children and are said to reflect systematic increases in memory span, speed of information processing, utterance length, and the use of memory strategies in encoding, storage and retrieval (Kail, 1990). Brainerd and Reyna (2004) provide a recent review of studies that have shown improvements in both verbatim and gist memory from early childhood through to young adulthood. Interest in the variability of verbatim and gist memory across development has emerged from studies of false memories; that is, memories consistent with the interpreted gist but not the verbatim facts (Brainerd, Reyna, & Forrest, 2002). Contrary to the notion that young children would be more vulnerable to false memories due to their limited verbatim memory, some examples of false memories have been found to occur more often in older children and adults than young children. As an explanation, the *fuzzy-trace* theory

(Reyna & Brainerd, 1995) proposes that false memories can occur through two independent routes; either decreases in verbatim memory or increases in gist processing. Some forms of gist processing have been found to be underdeveloped in young children (e.g., the spontaneous use of semantic clustering in recall; Bjorklund & Jacobs, 1985), and therefore reduced gist processing may lessen the formation of false memories early in development. Of relevance to the present study is the work by Brainerd and Mojardin (1998) who report developmental increases in the false recognition of gist-consistent sentences in children aged 6 to 11 years.

In general, adult females are reported to be superior to adult males on measures of verbal learning and memory (Lezak, 1995). Developmental studies have also found a female advantage on measures of verbal memory. Kramer, Delis, Kaplan, O'Donnell, and Prifitera (1997), for example, investigated gender differences on the California Verbal Learning Test in children between ages 5 and 16 years of age (N = 811). Females were found to recall more words than males on each list-learning trial and were more likely to use semantic organisation strategies to aid recall. This gender difference was apparent in each age group, with a trend for the size of the gender difference to increase with age.

Verbal memory in ASD

There is evidence to suggest that individuals with autism give preferential attention to the surface form of linguistic information rather than the meaning (see Chapter 2, Section 2.4.3). For example, individuals with autism have been found to have lower reading comprehension skills relative to their single-word reading ability (Frith & Snowling, 1983). Scheuffgen (1998) also found enhanced surface form retention in children and adolescents with autism compared to control participants matched on reading age. She observed that the individuals with autism were better able than controls to represent the exact wording of individual sentences presented in stories, but had more difficulty abstracting the overall meaning of these sentences.

Further evidence of a difficulty in using context when processing sentences has been reported by Jolliffe and Baron-Cohen (1999; 2000). In a series of studies, high-functioning adults with ASD were found to be impaired at integrating linguistic information in context, both at a local level of coherence (i.e., one to three sentences) and a more global level of coherence (i.e., five or more sentences). When required to rearrange sentences in order to make a coherent story, individuals with ASD were able to make use of temporal cues to link sentences but had greater difficulty when these cues were absent and global coherence was required to interpret the information. Jolliffe and Baron-Cohen also found that individuals with ASD showed a specific deficit in integrating the information to make inferences. Compared to control participants, the ASD group provided fewer correct explanations for the action of a story character when the reason needed to be deduced from the context.

As discussed above, a reduced tendency or ability to use context may result in individuals with autism being less prone to false memories. Beversdorf et al. (2000) tested this hypothesis using a false-memory paradigm whereby the effect of learning a list of related words typically induces incorrect recognition or recall for words not presented, but highly associated to the list (Roediger & McDermott, 1996). Beversdorf et al. found that high-functioning adults with ASD were more accurate in recognising words from the original list compared to control participants. In particular, the ASD adults were less likely to be fooled into falsely recognising associated words. Bowler et al. (2000) were not able to replicate this finding however, using the same false-memory paradigm.

Deficits in story recall have also been found in high-functioning adolescents and young adults with autism (Minshew & Goldstein, 2001), as measured by the number of story units recalled word for word in the Logical Memory subtest of the Wechsler Memory Scale-Revised (WMS R; Wechsler, 1987). It was proposed that the complexity of the information and the need for organisational strategies was challenging for the individuals with autism. In a recent study Williams et al. (2005) administered the latest revision of the Wechsler Memory Scale (WMS III; Wechsler, 1997b) to a group of high-functioning adolescents and adults with autism and age- and IQ-matched controls. In the updated Logical Memory subtest, scores for both the recall of details and for mentioning the main thematic units of the stories are assessed, but no significant group differences were found on this occasion. The authors surmised that as their participants were older and had higher VIQ than in their earlier study (Minshew & Goldstein, 2001), this might explain their inconsistent findings.

Predictions

Individual differences in verbal memory were assessed in the current study by means of reproduction and recognition. Participants were asked to read two short stories and immediately describe what each story was about. A recognition test followed to assess retention of surface form for individual sentences. Weak central coherence was predicted to be shown by enhanced surface form retention levels, together with decreased gist recall. Individuals with ASD are therefore expected to show greater ability to discriminate gist-consistent sentences from actually presented sentences than matched controls. A greater tendency for verbatim recall is also predicted in the ASD group. As developmental effects have been found in verbal memory, younger participants are suggested to show reduced verbatim and gist recall. However, according to studies of false memories, less influence of meaning may also result in greater retention of surface form in younger participants. As these speculations predict opposite effects, no specific predictions are made for developmental effects in verbatim memory. Gender differences in verbal learning have been demonstrated in the literature, and females are therefore predicted to show greater gist memory than males.

8.4.2 Method

Materials

Two stories (the Cave story and the Coin story) were selected for the present study from the set of stories created by Scheuffgen (1998) to assess surface form retention levels and representation of gist. Each story consisted of 11 sentences, with the Cave story consisting of 97 words and the Coin story consisting of 117 words (see Appendix N). The stories followed a schema described by Mandler (1978): (1) setting, (2) initiating event, conflict or change, and (3) resolution. Mental state attribution was not required to comprehend each story and all conflicts and changes described were due to events in the physical world. The stories had previously been used in a study with participants varying in age from 9 to 16 years (Scheuffgen, 1998) and were therefore considered age-appropriate for the present study. A short three-sentence story consisting of 37 words (the Football story) designed by the author was used as the practice item.

The alternative surface forms for the middle five sentences (i.e., sentences 4 to 8) as described in Scheuffgen (1998) were used in the present study. The new sentences were created by switching the order of clauses, using pronouns instead of nouns, and using synonyms (see Appendix N); the meaning of each sentence thereby remained unchanged. In the recognition test phase the alternative surface form was presented for three of the five sentences, while the remaining two sentences were presented in their original form. The higher number of sentences presented in a different form than identical form was deliberate in order to control for the high number of false positives expected with this task design (i.e., a "same" response bias).

Each story appeared as a paragraph with each sentence presented on a separate line, while the sentences used in the recognition test appeared individually. The story materials were printed on white A4 paper in portrait orientation and were presented to the participant in a test booklet.

Procedure

The participant was first made aware that they were to read (or listen to) some stories and answer some questions about what they had read (heard). To ensure the participant had adequate reading skills the participant was instructed to read the story aloud. To aid comprehension of the text the experimenter corrected all mispronunciations. The experimenter read the stories aloud if the participant was younger than 10 years or had low reading ability. In this situation, participants were encouraged to follow the story in the booklet as it was read to them.

As an introduction to the Story Memory Task the participant was instructed to read (or listen to) the practice story presented in the test booklet. On completion of the story, the participant was asked to read (or listen to) the story again. The test booklet was then turned to a blank page and the experimenter asked the participant: *"What was that story about?"* Participants had unlimited time to provide their description and one prompt was given (i.e., *"Can you tell me anything more about the story?"*) if a short description was given. All descriptions were audio recorded and later transcribed.

Immediately after the retelling of the story, the test sentences assessing surface form recognition were presented individually and in the same order as the original story. The experimenter read the sentence aloud and asked the participant *"Is this sentence exactly the same as it was in the story, or has it changed?"* Corrective feedback was provided on the practice story only and the experimenter turned back to the original story in order to emphasise how the test sentences could differ in form, while not differing in meaning. As previous studies have found that participants assume that they should respond on the basis of meaning unless explicitly stated (Reyna & Kiernan, 1994), surface form changes were clearly explained in order to avoid any misconceptions that responses should be based purely on gist.

After it was ensured that the participant understood the task requirements the Cave story was administered, followed by the Coin story. In each case, the story was read aloud twice and the reading time was unlimited. The participant was then asked to describe the events of the story, directly followed by the recognition test for surface form changes. The five test sentences (sentence 4 to 8) were presented individually and participants had unlimited time to make a same/different judgment to each. All responses were recorded by the experimenter on individual score sheets.

Qualitative data analysis

Participants' retellings of the two stories were scored for both global (gist) and local (detail) recall. As described in Scheuffgen (1998), the most important events were selected from each story for the scoring of gist. The Cave story was considered to consist of 13 story events, while the Coin story divided into 11 story events (see Appendix O). Each story event received a score of 1-point if the gist of the event was captured in the participant's description. Details such as exact names and places or illustrative adjectives were not required. If the general idea of the event was not mentioned or was incorrectly interpreted, the story event scored 0-points. The maximum number of story events obtainable across both stories was 24.

The scoring system for local detail recall was devised for the present study by the author. Each story was divided into units of distinct information: 27 units were identified in the Cave story and 34 units in the Coin story (see Appendix O). Each story unit obtained a score of 1-point if the participant had mentioned the detail in their description. Allowance was made for variations of verbs (e.g., spotted / found / saw), although correct proper names were necessary for scoring. The correct order of recall of the story units was not a requirement, but a single word or phrase used by the participant could not be scored more than once. The maximum *local detail* score across both stories was 61.

A more strict measure of local detail recall was also taken whereby participants were scored for recalling a phrase from the original story word-for-word. A phrase was defined as two or more words that formed an expression. The number of distinctive phrase units was limited to 24 in the Cave story and to 30 in the Coin story (see Appendix O). A score of 1-point was assigned if the participant recalled a phrase verbatim within the defined unit. For example, for the phrase unit *"when they woke up"* a score was assigned if the participant had recalled, *"when they woke"* or *"they woke up."* In order to obtain a score that reflected style of memory, rather than amount of recall, the phrase recall score was expressed as a proportion of the total number of phrase units recalled (irrespective of style of recall) by the individual for each story. The proportion of verbatim recall was averaged across the two stories for each participant.

The inter-rater reliability of the scoring system for local and global descriptions is presented in Appendix O.

8.4.3 Typical Development Results

All 204 TD participants were administered the Story Memory Task. Data were present from all participants for both the Surface Form Recognition Test and the Free Recall of the story. The analyses were performed with and without participants who had ESL. As pattern of findings was not altered, data from participants with ESL remained in the analysis.

Story Description

The mean number (and percentage) of gist elements and local detail units recalled by each age group is presented in Table 8.5. Also shown is the mean proportion of phrase units recalled across the two stories. As this latter variable was normally distributed in each age group, parametric statistical analyses were used. The data for the total number of gist elements and local detail units were negatively skewed in each age group (z-scores ranged from –1.0 to –4.0), and non-parametric statistical tests were therefore used for these variables.

No significant effects of age group were found for the number of gist elements ($\chi^2 = 4.31$, $p = .23$) or local detail units recalled ($\chi^2 = 2.07$, $p = .56$). Furthermore, no age group effects were found when analysing male and female participants separately (all $\chi^2 < 5.85$, $p > .12$). Male and female participants also did not differ overall on the mean number of gist elements ($z = 0.36$, $p = .72$) or local detail units recalled ($z = 0.19$, $p = .85$).

Table 8.5 Mean number (and *percentage*) of gist elements and local detail units recalled and proportion verbatim phrase recall on the Story Memory Task for each age group: Mean (*SD*).

Age group (years)	Gist elements (max = 24)		Local detail units (max = 61)		Proportion of verbatim phrase recall
8-10	17.0 (4.3)	*71%*	35.3 (9.8)	*58%*	0.52 (0.14)
11-13	17.0 (4.5)	*71%*	35.4 (11.2)	*58%*	0.53 (0.13)
14-16	17.4 (5.2)	*73%*	35.6 (13.0)	*58%*	0.52 (0.13)
17-25	15.9 (5.1)	*66%*	32.8 (12.6)	*54%*	0.49 (0.15)

A two-way ANOVA performed on the proportion of verbatim phrase recall did not show a significant main effect of age group ($F_{(3,196)} = 0.89$, $p = .45$), gender ($F_{(1,196)} = .29$, $p = .59$), or interaction between age group and gender ($F_{(3,196)} = 0.38$, $p = .77$). Adjusting for the effects of FIQ (or FIQ-BD) did not alter the pattern of findings. As shown in Figure 8.7 only a slight decrease in verbatim phrase recall is evident with increasing age in both male and female participants.

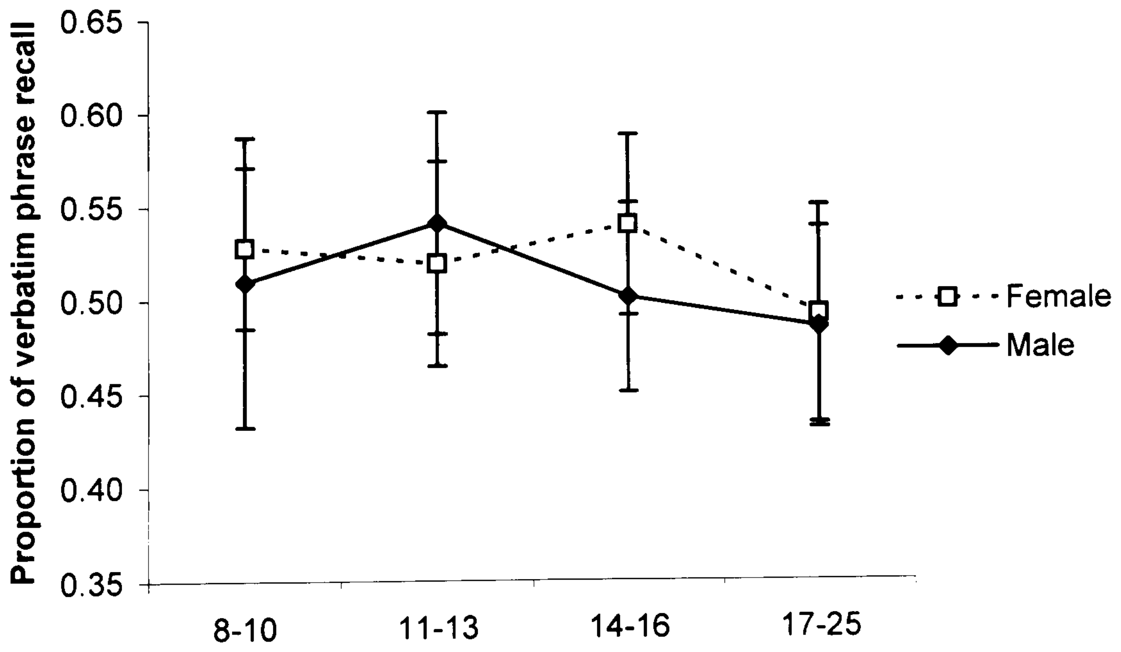


Figure 8.7 Proportion of verbatim phrase recall for male and female participants in each age group. Error bars show 95% confidence intervals.

Correlations: Story Description

The proportion of verbatim phrase recall showed a negative association with age in the TD sample, although the correlation did not reach significance ($r = -.12$, $p = .08$; females $r = -.14$, $p = .14$; males $r = -.09$, $p = .36$). A positive correlation was found between the proportion of verbatim phrase recall and IQ across all TD participants (FIQ $r = .18$, $p = .008$; VIQ $r = .18$, $p = .01$, PIQ $r = .13$, $p = .06$, FIQ-BD $r = .20$, $p = .005$). A similar pattern was observed in male participants (FIQ $r = .20$, $p = .05$; VIQ $r = .23$, $p = .02$, PIQ $r = .10$, $p = .32$, FIQ-BD $r = .24$, $p = .02$). No correlations between the proportion of verbatim phrase recall and the IQ measures reached significance in female participants (FIQ $r = .18$, $p = .07$; VIQ $r = .12$, $p = .23$, PIQ $r = .18$, $p = .07$, FIQ-BD $r = .15$, $p = .12$), although the size of the correlation coefficients did not differ between genders (all $z_{r1\text{-}r2} < 0.85$, $p > .40$).

The proportion of verbatim phrase recall showed a positive correlation with the number of gist elements recalled ($r_s = .40$, $p < .0005$). This was evident in male participants ($r_s = .54$, $p < .0005$) and female participants ($r_s = .27$, $p = .006$), and suggests that verbatim memories and gist are positively dependent.

Surface Form Recognition Test

The number (and percentage) of correct same/different judgments in the Surface Form Recognition Test and the corresponding measure of sensitivity A' are presented in Table 8.6 for each age group. As the data failed tests of normality (all $p < .01$), non-parametric statistical analyses were used.

No significant effects of age group were found on the number of correct judgments of *same* sentences for all TD participants ($\chi^2 = 2.87$, $p = .41$), and for each gender (male $\chi^2 = 2.75$, $p = .43$; female $\chi^2 = 3.07$, $p = .38$). Significant age group effects were however found on the number of correctly detected *different* sentences ($\chi^2 = 11.39$, $p = .01$). Pairwise comparisons revealed that the youngest age group judged significantly fewer sentences correctly that had changed in surface form than all older age groups (all $z > 2.06$, $p < .02$). Furthermore, the 8-10 age group did not differ from chance levels in correctly judging *different* sentences ($t_{(53)} = 1.51$, $p = .14$); whereas all older age groups were able to detect *different* sentences at significantly greater than chance levels (all $t > 5.80$, $p < .0005$). The significant effect of age group on correct detections of *different* sentences was found in male participants ($\chi^2 = 9.83$, $p = .02$), but not in female participants ($\chi^2 = 5.65$, $p = .13$). As shown Figure 8.8, male participants in the youngest age group performed significantly lower in identifying sentences that had changed in surface form than all older age groups (all $z > 1.88$, $p < .03$).

A main effect of gender was found on *different* form detections, with females identifying significantly more sentences that were altered from the original than males ($z = 2.25$, $p = .03$). When comparing within age group, this gender effect was only apparent in the youngest age group ($z = 2.19$, $p = .03$). No gender effects were found on the number of correctly judged *same* sentences ($z = 1.01$, $p = .31$).

Table 8.6 Number (and *percent*) correct and A' for same/different judgments of test sentences on the Surface Form Recognition Test for each age group: Mean (*SD*).

Age group (years)	Same sentences (max = 4)		Different sentences[a] (max = 6)		A'
8 - 10	3.06 (0.83)	*76%*	3.28 (1.35)	*55%*	0.72 (0.21)
11 - 13	3.19 (0.73)	*80%*	3.95 (1.02)	*66%*	0.81 (0.13)
14 - 16	3.14 (0.66)	*78%*	4.04 (1.22)	*67%*	0.80 (0.14)
17 - 25	2.89 (0.87)	*72%*	3.86 (1.10)	*64%*	0.76 (0.16)

[a]8-10 < 11-13, 14-16, 17-25, all $p < .02$.

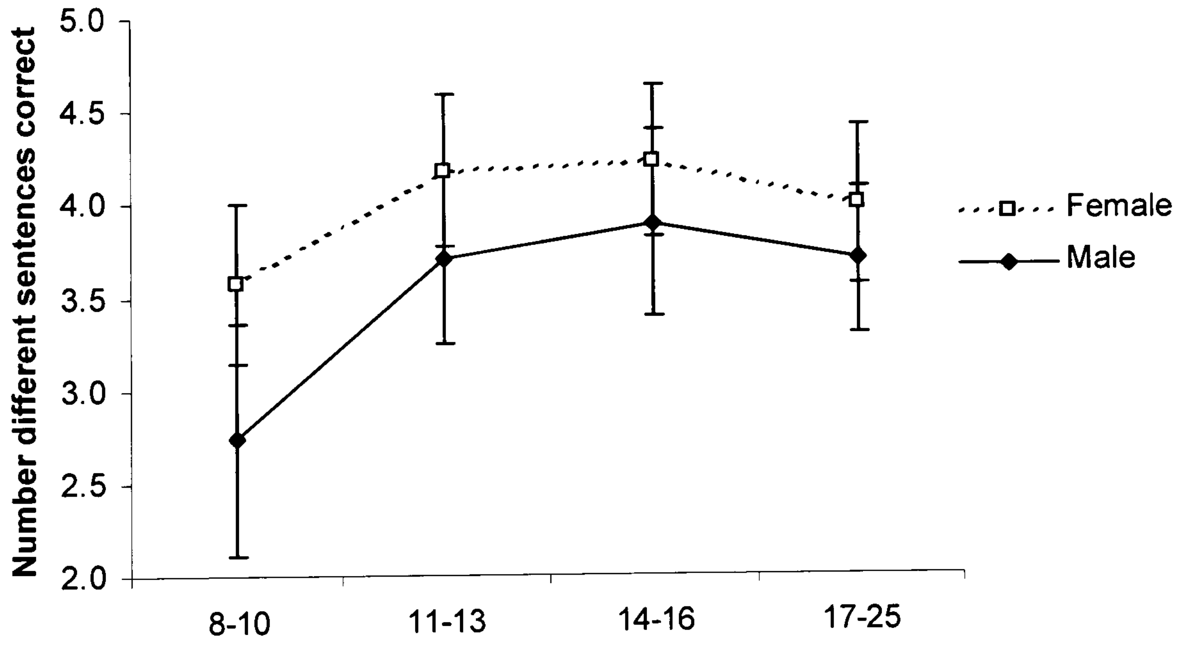


Figure 8.8 Mean number of 'different' sentences judged correctly (max = 6) for male and female participants in each age group. Error bars show 95% confidence intervals.

Signal detection analyses were conducted in order to see whether the age group differences could be accounted for by a response bias, B''. Although the youngest age group had a negative B'' ($M = -.42$, $SD = .59$), indicating a tendency to respond "same" to all sentences, no significant effects of age group were found on this measure overall, or within each gender (all $\chi^2 < 3.99$, $p > .25$). Male participants were found to have a significantly lower B'' ($M = -.41$, $SD = .59$) than females ($M = -.23$, $SD = .64$; $z = 2.09$, $p = .04$), suggesting males were more likely to respond "same" on each trial. When analysing by age group, the male and female difference failed to reach significance (all $z < 1.69, p > .08$), although B" was more negative for males than females in all age groups.

No significant effect of age group was found on the sensitivity measure A' which takes response bias into account ($\chi^2 = 5.93$, $p = .12$), although a trend was apparent in male participants ($\chi^2 = 6.65$, $p = .08$) which was absent in female participants ($\chi^2 = 2.64$, $p = .45$). Male and female participants did not differ on A' overall, ($z = 0.24$, $p = .81$), or within each age group (all $z < 1.59, p > .10$).

Correlations: Surface Form Recognition

The sensitivity measure A' did not correlate with age in the TD sample overall ($r_s = .04$, $p = .53$), nor within male ($r_s = .10$, $p = .32$) or female participants ($r_s = -.02$, $p = .85$; $z_{r1\text{-}r2} = 0.85$, $p = .40$). The only IQ measure found to correlate with test performance was PIQ ($r_s = .14, p = .05$); no other IQ index reached significance ($r_s = .05$ to $.10$, all $p > .16$). Within TD males a significant positive correlation was found between A' and PIQ ($r_s = .20, p = .05$), and FIQ-BD ($r_s = .20, p = .05$), but did not reach significance for FIQ ($r_s = .17, p = .09$), nor VIQ ($r_s = .15, p = .14$). No association was found between IQ and A' within TD females ($r_s = -.10$ to $.07$, all $p > .44$). No significant difference was found in the strength of the correlation coefficients between genders (all $z_{r1\text{-}r2} < 1.53$, $p > .12$).

No association was found between the two measures of local processing taken from the Story Memory Task. Surface form detection (A') and proportion of verbatim phrase recall showed no signification correlation ($r_s = .002, p = .97$; males $r_s = .03, p = .79$; females $r_s = -.04$, $p = .72$), even after controlling for the effects of age and FIQ ($pr = .01, p = .89$; males $pr = -.02, p = .85$; females $pr = .04, p = .72$).

8.4.4 *ASD and Control Results*

The Story Memory Task was not administered to one 14-year-old ASD participant (FIQ = 49) and one 16-year-old control participant (FIQ = 60). Both participants had low reading ability and were also not able to follow the story along as it was read aloud by the experimenter. In addition, one 14-year-old from the ASD group who had not attended the final test session was not administered the Story Memory Task. The groups remained well matched on age and all IQ measures with the removal of these three participants from analyses (all $t_{(56)} < 0.17, p > .86$).

Story Description

Table 8.7 presents the mean number (and percentage) of gist elements and local detail units recalled for each group. The proportion of verbatim phrase recall for the two stories is also presented. As found in the TD sample, the mean number of gist elements and local detail units data failed tests of normality, although the proportion of verbatim phrase recall data were normally distributed for both groups. Non-parametric statistical analyses were therefore conducted for the gist and local detail recall data, while parametric statistics were used on the proportion of verbatim phrase recall data.

Significant group differences (all two-tailed tests) were found for the number of gist elements ($z = 3.45$, $p = .001$) and local detail units recalled ($z = 3.22$, $p = .001$). In both cases the ASD group recalled significantly fewer aspects relative to the control group. The same pattern of results was found when only including high-functioning participants in the group comparisons (gist elements $z = 3.20$, $p = .001$; local detail units $z = 3.09$, $p = .002$). When comparing low-functioning participants, however, the differences in the amount of gist and detail recall did not significantly differ (gist elements $z = 1.47$, $p = .14$; local detail units $z = 0.74$, $p = .46$).

Table 8.7 Mean number (and *percentage*) of gist elements and local detail units recalled and proportion verbatim phrase recall on the Story Memory Task for each group overall, and split by high and low IQ: Mean (*SD*).

	Group	N	Gist elements[a] (max = 24)		Local detail units[b] (max = 61)		Proportion of verbatim phrase recall[c]
All	ASD	29	9.9 (5.9)	*41%*	18.7 (11.6)	*31%*	0.39 (0.21)
	Control	30	15.7 (5.9)	*65%*	30.5 (13.7)	*50%*	0.51 (0.16)
High IQ	ASD	24	10.8 (6.0)	*45%*	20.2 (12.0)	*33%*	0.36 (0.21)
	Control	25	16.5 (5.2)	*69%*	32.7 (13.0)	*54%*	0.52 (0.14)
Low IQ	ASD	5	5.4 (3.7)	*23%*	11.2 (5.8)	*18%*	0.56 (0.14)
	Control	5	11.6 (7.6)	*48%*	19.6 (13.1)	*32%*	0.47 (0.24)

[a] ASD < Control, $p = .001$; ASD_{high} < $Control_{high}$, $p = .001$.
[b] ASD < Control, $p = .001$; ASD_{high} < $Control_{high}$, $p = .002$.
[c] ASD < Control, $p = .01$; ASD_{high} < $Control_{high}$, $p = .002$.

As predictions made from weak central coherence theory related to style of recall, rather than absolute amount, one-tailed t-tests were performed to compare groups on the mean proportion of verbatim phrase recall. A significant group difference was found on this variable ($t_{(57)} = 2.41$, $p = .01$). The finding was opposite to predictions however, as the control group produced relatively more word-for-word recall than the ASD group. This was also found when comparing groups of participants with high IQ ($t_{(40.25)} = 3.06$, $p = .002$), while no group differences were observed between groups of participants with low IQ ($z = 0.63$, $p = .13$). It is worth noting however, that the mean proportion of verbatim phrase recall for the low IQ ASD participants was higher than the low IQ control participants, and higher than any age group in the TD sample (see Table 8.5).

Correlations: Story Description

The proportion of verbatim phrase recall was negatively correlated with age in the control group, although not at a significant level ($r = -.28$, $p = .13$). No significant association between the two variables was found in the ASD group ($r = -.03$, $p = .86$; $z_{r1-r2} = 0.93$, $p = .36$). The proportion of verbatim phrase recall did not correlate with MA in either group (ASD $r = -.16$, $p = .42$; control $r = .08$, $p = .67$). Positive correlations were found between IQ and the proportion of verbatim phrase recall in the control group, which reached significance for VIQ ($r = .38$, $p = .04$) and FIQ-BD ($r = .36$, $p = .05$), but not for FIQ ($r = .30$, $p = .10$) or PIQ ($r = .13$, $p = .51$). In contrast, negative correlations were found between the IQ measures and the proportion of verbatim phrase recall in the ASD group, none of which approached significance ($r = -17$ to $-.10$; all $p > .37$). Furthermore, the strength of correlation between VIQ and verbatim phrase recall significantly differed for ASD and control participants ($z_{r1-r2} = 2.08$, $p = .04$).

A positive correlation was found in both groups between the proportion of verbatim phrase recall and the total number of gist elements recalled. This association was significant in the control group ($r_s = .40$, $p = .03$) and approached significance in the ASD group ($r_s = .33$, $p = .08$; $z_{r1-r2} = 0.31$, $p = .76$).

Surface Form Recognition Test

Table 8.8 presents the mean number (and percent) correct same/different judgments and the corresponding measure of sensitivity A' for each group on the Surface Form Recognition Test. As each variable failed to pass tests of normality within each group, non-parametric statistics were used.

The ASD group performed significantly above chance in correctly detecting *same* sentences ($t_{(28)} = 10.21$, $p < .0005$) and *different* sentences ($t_{(28)} = 3.08$, $p = .005$). In contrast, the control group performed significantly above chance on *same* sentences ($t_{(29)} = 8.84$, $p < .0005$), but not *different* sentences ($t_{(29)} = 1.88$, $p = .07$).

One-tailed Mann-Whitney U tests found no significant group differences for the mean number of correctly judged *same* sentences ($z = 0.54$, $p = .29$) or *different* sentences ($z = 1.03$, $p = .15$). No group differences were found with high-functioning participants for *same* sentences ($z = 0.07$, $p = .47$) and *different* sentences ($z = 0.28$, $p = .39$). Although the number of participants was low (N = 5 in each group), low IQ participants in the ASD group correctly determined sentences that differed from the original significantly more often than low IQ control participants ($z = 1.84$, $p = .03$). No difference was observed between the low IQ groups in correctly detecting unchanged sentences ($z = 1.32$, $p = .09$).

It was notable that three participants from the ASD group had perfect scores in determining whether the test sentence had changed from the original story (aged 11.2 years, FIQ = 122; 13.8 years, FIQ = 78; 15.8 years, FIQ = 84). Perfect performance was never observed in the control group, and only once in the TD sample (female aged 11.5 years, FIQ = 110).

Table 8.8 Number (and *percent*) correct and A' for same/different judgments of test sentences on the Surface Form Recognition Test for each group overall, and split by high and low IQ: Mean (*SD*).

	Group	N	Same sentences (max = 4)		Different sentences[a] (max = 6)		A'[a]
All	ASD	29	3.38 (0.73)	*84%*	3.90 (1.57)	*65%*	0.82 (0.19)
	Control	30	3.27 (0.78)	*82%*	3.50 (1.46)	*58%*	0.78 (0.17)
High IQ	ASD	24	3.29 (0.75)	*82%*	3.96 (1.57)	*66%*	0.81 (0.18)
	Control	25	3.32 (0.69)	*83%*	3.88 (1.24)	*65%*	0.82 (0.13)
Low IQ	ASD	5	3.80 (0.45)	*95%*	3.60 (1.67)	*60%*	0.82 (0.24)
	Control	5	3.00 (1.22)	*75%*	1.60 (0.89)	*27%*	0.57 (0.23)

[a] $\text{ASD}_{\text{low}} > \text{Control}_{\text{low}}$, $p = .03$

Signal detection analyses were conducted in order to take into account any differences in response bias between the two groups. Both groups showed a strong bias towards saying "same" when determining whether a test sentence had altered in form (ASD $M = -.47$, $SD = .67$; control $M = -.48$, $SD = .64$). No significant group difference was observed on B'' ($z = .03$, $p = .97$, two-tailed; high IQ $z = 0.54$, $p = .59$; low IQ $z = 1.42$, $p = .16$).

No group difference was found on the sensitivity measure A' overall ($z = 1.27$, $p = .10$, one-tailed), nor in the high-functioning participants ($z = 0.25$, $p = .40$). When comparing the group performance of low-functioning participants, the ASD group showed superior ability to detect sentences presented in different form ($z = 1.90$, $p = .03$).

Correlations: Surface Form Recognition

No reliable correlation was observed between A' and age in the ASD group ($r_s = .02$, $p = .91$) or the control group ($r_s = -.31$, $p = .10$; $z_{r1\text{-}r2} = 1.20$, $p = .23$). Furthermore, no significant correlation was found between MA and A' in either group (ASD $r_s = -.19$, $p = .33$; control $r_s = .16$, $p = .41$; $z_{r1\text{-}r2} = 1.29$, $p = .20$). Significant correlations were found between IQ and A' in the control group (FIQ $r_s = .51$, $p = .004$; FIQ-BD $r_s = .49$, $p = .01$; PIQ $r_s = .54$, $p = .002$; VIQ $r_s = .38$, $p = .04$). In contrast, no association was found between IQ and A' in the ASD group (r_s ranged from -.18 to -.10, all $p > .33$). The size of the correlation coefficients between the IQ indices and A' were significantly different between groups (all $z_{r1-r2} > 2.11, p < .04$).

No association was found between verbatim phrase recall and surface form detection (A') in the ASD group ($r_s = .14, p = .48$) or the control group ($r_s = .02, p = .94$), even after controlling for the effects of age and FIQ (ASD $pr = .05$, $p = .80$; control $pr = -.34$, $p = .08$).

8.4.5 Summary and brief discussion of Story Memory Task

Although overall ability to discriminate gist-consistent sentences from actual sentences did not vary with age in the TD sample, age effects were found (primarily in male participants) for the correct detection of sentences that had differed from the original story. Female participants showed greater proficiency than male participants to correctly identify sentences that had changed, with this gender difference being significant in the youngest age group. This finding is contrary to the study by Kramer et al. (1997) who found the magnitude of the gender difference in verbal memory (recalling word lists) to increase with age. The role of PIQ was evident in the ability to discriminate between gist-consistent sentences and original sentences; again, this effect was more apparent in male than female participants.

The absolute number of gist elements and local detail units recalled by participants did not differ with age or between genders in the TD sample. A trend was shown for younger participants to recall the story more word-for-word than older participants, although no significant variation with age was found. Male and female participants also did not differ in their style of recall. IQ was associated with the proportion of verbatim phrase recall, indicating that participants with higher IQ tended to recall the story word-for-word.

No group differences were found in discriminating between gist-consistent sentences and original sentences, however the ASD group were able to identify sentences that had changed in surface form better than chance performance, while the control group performed at chance levels. Furthermore, ASD participants with low IQ were more proficient than their low-functioning counterparts at determining a sentence had changed in surface form, performing as well as high-IQ ASD participants and TD adolescents.

The ASD group recalled significantly fewer gist elements and local detail units than the control group. This replicates Minshew and Goldstein (2001) who found lower recall for story units in high-functioning individuals with autism. It also confirms the work of Bowler and colleagues (e.g., Bowler, Matthews, & Gardiner, 1997), who found that individuals with ASD typically show lower recall than recognition ability. Contrary to predictions of weak central coherence however, the ASD group did not show a tendency to recall the stories verbatim. Both recognition ability and tendency to recall verbatim were determined by general intelligence in the control group, but showed no relationship with ability in the ASD group.

It was predicted that weak central coherence would be evidenced by a greater retention of surface form along with reduced gist memory. However both processes were found to be positively associated in the TD sample, as well as in the control and ASD groups. Gist recall was therefore not compromised by good retention of surface form, as suggested by the literature on false memories. The current finding is consistent however with the work of Gernsbacher (1985), who proposed that there is not necessarily a trade-off between surface form retention and comprehension abilities; verbatim memory can also be a consequence of good comprehension.

A major criticism of the current task is the simultaneous assessment of gist memory and surface form retention. Participants were explicitly instructed to retain surface information, and were tested on this level of information following the free recall test. It is possible that the conscious effort made by participants to retain surface information may have influenced the free recall phase, resulting in the tendency to recall each story word-for-word. Participants were made aware that they were going to be tested for memory of surface information in order to avoid any misunderstandings that responses should be made on the basis of meaning. However Murphy and Shapiro (1994) maintain that verbatim memory should be studied as naturalistically as possible and participants should not be told to memorise the text. They argue that memorisation processes are not the same as comprehension processes; if the participant is aware of the verbatim memory test, the results are likely to be different from normal comprehension and will preclude the study of normal language processing. Future studies should therefore separate the deliberate measurement of surface form retention from the assessment of style of recall. An open-ended approach should be taken in order to assess an individual's natural tendency for gist versus verbatim memory of normal discourse.

Two candidate variables were taken from the Story Memory Task as continuous measures of performance: sensitivity of discrimination of surface form changes (A') and the proportion of verbatim phrase recall. High scores on both measures are suggested to reflect weak central coherence in the processing of verbal material. Participants were categorised as showing weak central coherence if they were well able to discriminate between gist-consistent and actual sentences, as determined by A' values above .80.

8.5 General Discussion

The purpose of the experimental tasks reported in Chapter 8 was to determine the extent of individual differences in verbal coherence and to test the hypothesis that individuals with ASD would have difficulty deriving meaning from verbal information. A summary of task findings is presented in Table 8.9 for the TD sample and in Table 8.10 for the ASD and control group comparisons (with effect sizes presented in Appendix Q). Table 8.11 presents the percentage of participants in each group characterised as showing weak central coherence on each verbal measure.

Developmental effects were found in the TD sample with older participants showing greater global processing and sensitivity to context. This was demonstrated by providing coherent completions to sentence stems, correcting errors in the reading of homographs and, to some extent, showing less tendency to recall a story verbatim. Position effects in the reading of homographs in context, predicted to be indicative of global processing, were less apparent in older participants however. This finding may have been as a result of ceiling effects and a greater tendency to read ahead in older experienced readers. IQ was also related to task performance; individuals with high IQ were more likely to correct errors spontaneously when reading homographs in context, and retain verbatim information from short stories.

Some gender effects were found on the verbal coherence measures. Younger male participants were more likely to provide local completions to sentence stems, suggesting a local processing bias in males. However, male participants also showed a greater effect of position when reading homographs in context, which was predictive of global processing bias. IQ determined the extent of position effects in male participants, such that those with lower IQ showed a greater effect of context position. Male participants with low IQ also showed lower retention of surface form information, which indicates that IQ may drive local processing in male participants.

Although developmental and gender effects on task performance were found on the verbal coherence measures, individual differences over and above these effects were found in the TD sample. As shown in Table 8.9, a significant proportion of the variance from each coherence index could not be explained by age, IQ, and gender and may reflect differences in cognitive style.

It was predicted that, compared to control participants, individuals with ASD would be less proficient at completing a sentence meaningfully, show reduced use of context to disambiguate homographs, and show greater retention of surface-level than semantic-level information when reading stories. Predictions were supported on the Sentence Completion Task, with the frequency of local completions and/or long response times found to be significantly higher in the ASD group compared to the control group. This local response style was not evident in ASD participants with low IQ however. As predicted, ASD participants were less likely to correct errors made when reading homographs in context, although overall sensitivity to sentence context, as shown an effect of homograph position, was not demonstrated in the ASD group. Correct recognition of surface-level information was above chance levels in the ASD group, while only at chance in the control group. However a verbatim style of recall was not evidenced in the ASD group, and was in fact significantly more common in the control group. A significant reliance on IQ across the coherence measures was found in the control group as in the TD sample, whereas no association between IQ and task performance was found in the ASD group. As the tendency to recall stories verbatim appeared to tap general intelligence in the control group, rather than response style, this may explain why the expected group difference was not found on this measure.

Predictions from weak central coherence were most clearly supported on the Sentence Completion Task. In comparison to the passive reading of sentences and short stories, the Sentence Completion Task required an immediate response from the participant. Weak central coherence may therefore have been better captured by this open-ended response format. Although the Homograph Reading Task was designed to be as open-ended as possible, ASD participants may have had an expectation to read for meaning. This is a typical requirement when asked to read aloud in a school setting and therefore an individual's natural reading style may not have been measured.

The inter-relationships between the three verbal coherence tasks presented here are explored in Chapter 9. In addition, the associations of these high-level verbal measures with coherence measures from different processing modalities are investigated.

Table 8.9 Summary of high-level verbal task findings in the TD sample.

Verbal Task	Dependent Variable (Direction of high values)	Age effects?	IQ effects?	Gender effects?	Proportion of variance not explained by age, IQ and gender	Interactions
Sentence Completion[a]	Error Score (Weak CC)	Yes, negative	No	M > F (trend)	95%	Negative correlation with age in M only.
Homograph Reading	Effect of position (Strong CC)	Yes, negative	No	M > F	94%	Negative correlation with IQ in M only.
	Proportion of self-corrections (Strong CC)	Yes, positive	Yes, positive	No	90%	F > M at 14-16
Story Memory	Proportion of verbatim phrase recall (Weak CC)	No (negative trend)	Yes, positive	No	96%	
	Surface form retention, A′ (Weak CC)	No	Yes, positive PIQ	No	99%	Positive correlation with PIQ in M only

[a]N = 192 with English as a first language. Note: M = males, F = females

Table 8.10 Summary of high-level verbal task findings in the ASD and control groups.

Verbal Task	Dependent Variable (Direction of high values)	Group effects?	Age effects?	MA effects?	IQ effects?	Other comments
Sentence Completion	Error Score (Weak CC)	ASD > Control	No	No	Negative in ASD	ASD_{high} > $Control_{high}$
Homograph Reading	Effect of position (Strong CC)	No	No	No	Negative in Control	
	Proportion of self-corrections (Strong CC)	ASD < Control	No	Yes, positive	Positive in Control	ASD_{high} < $Control_{high}$
Story Memory	Proportion of verbatim phrase recall (Weak CC)	ASD < Control (not as predicted)	No	No	Positive VIQ/FIQ-BD in Control	ASD_{high} < $Control_{high}$ (not as predicted)
	Surface form retention, A′ (Weak CC)	No	No	No	Positive in Control	ASD_{low} > $Control_{low}$

Table 8.11 Percentage (and N/total N) of participants in each group who showed weak central coherence on each task.

Verbal Task	Marker for weak CC	8-10	11-13	14-16	17-25	ASD	Control
Sentence Completion[a]	Two or more local completions	17% (9/54)	23% (9/40)	12% (6/49)	10% (5/49)	39% (12/31)	10% (3/30)
Homograph Reading[b]	Two or more errors on rare homographs	37% (20/54)	30% (13/43)	12% (6/50)	23% (12/53)	43% (13/30)	21% (6/28)
Story Memory	Proportion of verbatim phrase recall $\geq .55$[c]	44% (24/54)	40% (17/43)	39% (20/51)	38% (21/56)	14% (4/29)	37% (11/30)
	$A' > .80$[d]	39% (21/54)	47% (20/43)	53% (27/51)	34% (19/56)	66% (19/29)	47% (14/30)

[a]ASD > Control, $p = .005$
[b]8-10 > 14-16, $p = .06$; ASD > Control, $p = .04$
[c]ASD < Control, $p = .02$
[d]ASD > Control, $p = .07$

Chapter 9. Individual differences in central coherence across tasks

9.1 Introduction to data reduction procedure

The main focus of this thesis was to determine whether pervasive individual differences could be found in central coherence, across modality and level of processing. To address this question, it was first determined whether measures classified into the same processing-level/modality quadrant were tapping a common underlying mechanism of coherence. If affirmed, composite scores were constructed to increase the reliability and power of the findings by minimising the number of statistical comparisons. Task clustering/composites were based on patterns of data in the large TD sample, which then determined the data reduction approach in the ASD and control groups.

A summary of the key coherence indices from each task in the coherence battery is presented in Table 9.1. Correlational analyses were first used to explore the relationship between coherence indices within each processing-level/domain. To facilitate comparisons, all coherence indices were coded so that a positive score reflected strong coherence (global bias), and positive correlations could therefore be predicted between tasks. Coherence indices that were reversed are indicated in Table 9.1. On occasion more than one candidate variable from a task was selected as an index of coherence. For reasons of brevity, the variable that showed greater inter-task correlation within the processing-level/domain is reported. The redundant coherence indices are shown in italics in Table 9.1.

Where appropriate, data reduction techniques (i.e., principal component analysis, PCA) were employed to detect structure in the relationships between variables. Kaiser-Meyer-Olkin's (KMO) test for measuring sampling adequacy and Bartlett's test of sphericity were conducted to assess the suitability of the data for factor analysis. The KMO index indicates the proportion of variance in the variables that is common variance and may reflect underlying factors. This index ranges from 0 to 1, reaching 1 when each variable is perfectly predicted without error by the other variables. The Bartlett test of sphericity tests for the presence of correlations among the variables. It determines whether the correlation matrix is an identity matrix, which indicates that the variables are unrelated and not suitable for factor analysis.

Prior to entering the data into the PCA, the confounding effects of age and ability (FIQ-BD) were removed by linear regression, using the saved standardised residuals as the adjusted scores. This approach was also conducted on ASD and control group data; in each case the removal of age and ability was based on their respective group data. This is a conservative approach, looking for consistent individual differences in central coherence, independent of the effects of age or IQ.

The data reduction analyses for each processing-level/modality quadrant are presented in the following sections, for the TD sample and the ASD and control groups. An examination of the consistency of coherence processing style across low- and high-level tasks and across visuo-spatial and auditory-verbal modalities is then provided.

Table 9.1 Coherence tasks and main indices, divided by processing-level/domain.

Domain	Task	Coherence Index[a]
Low-level visuo-spatial	Un/segmented Block Design	Relative benefit from segmentation
	Embedded Figures Test	Detection time
	Impossible Figures	A' for discrimination im/possible figures
	Navon Similarity Judgment	Total global matches
Low-level auditory	Phoneme Segmentation[R]	A' for detection of phoneme in non-words *Relative effect of position*
	Chord Segmentation[R]	Correct detections of target note in a chord *A' for discrimination of a note within a chord*
	Pitch Identification[R]	Total tone-animal pairs correctly identified
	Chord Sequence Task	Number GRLU target chords judged correct *Number GULR target chords judged correct*[R]
High-level visuo-spatial	Drawing Style (Rey & Pram)	Coherence composite for copy (order & style) *Effect of meaning on recall accuracy (Pram minus Rey)*
	Picture Memory: Description	Global description composite
	Picture Memory: Recognition[R]	A' for discrimination of surface form changes
	Fragmented Pictures[R]	Response time for correct identification
High-level verbal	Sentence Completion[R]	Local completions error score
	Homograph Reading	Effect of position (all words, initial response) *Proportion of self-corrections*
	Story Memory: Recall[R]	Proportion of verbatim recall
	Story Memory: Recognition[R]	A' for discrimination of surface form changes

[R]Score reversed in order for positive scores to reflect strong global coherence.
[a]Coherence indices in italics not reported due to overlap with task indices with higher correlation within domain/level.

9.2 Low-level visuo-spatial domain

Typical Development Results

Table 9.2 presents the inter-correlations among the four visual-perceptual measures of coherence in the TD sample. Zero-order parametric correlation coefficients are presented, along with partial correlation coefficients to show how removing the effects of age, FIQ-BD, and both age and FIQ-BD affected the relationships between tasks.

Table 9.2 Correlation matrix for the low-level visuo-spatial coherence measures in the TD sample.

		Un/segmented Block Design	Impossible Figures	Navon Similarity Judgment
Embedded Figures Test	***r***	**.46****	**-.34****	**-.25****
	pr_{age}	.38**	-.29**	-.14*
	$pr_{FIQ\text{-}BD}$	.38**	-.23**	-.21**
	$pr_{age\ \&\ FIQ\text{-}BD}$	.28**	-.16*	-.08
Un/segmented Block Design	***r***		**-.38****	**-.14~**
	pr_{age}		-.35**	-.03
	$pr_{FIQ\text{-}BD}$		-.29**	-.09
	$pr_{age\ \&\ FIQ\text{-}BD}$		-.24**	.03
Impossible Figures	***r***			**.23****
	pr_{age}			.18*
	$pr_{FIQ\text{-}BD}$			.18**
	$pr_{age\ \&\ FIQ\text{-}BD}$			.12~

~ p <.10, *p < .05 **p < .01

The factorability of the data (standardised residual scores) was affirmed by Bartlett's test of sphericity (p < .0005). However, the KMO index was .57 indicating that the degree of common variance among the four variables was classified as "miserable" bordering on "mediocre" according to the criteria set by Kaiser (1974). The factors extracted will therefore account for a fare but not a substantial amount of variance. With this in mind, the data were subjected to PCA using varimax orthogonal rotation[9]. The analysis yielded two factors with an eigenvalue larger than 1.00, which together explained 63% of the variance in the data. For clarity, all rotated factor loadings greater than .40 are presented in Table 9.3.

[9] Nonorthogonal rotation methods were also used (e.g., direct oblimin) but yielded the same result suggesting that the underlying factors are orthogonal.

Table 9.3 Factor loadings (> .40) from the PCA (orthogonal rotation) of the four low-level visuo-spatial coherence tasks in the TD sample with age and FIQ-BD removed.

Low-level visuo-spatial tasks	Factor Loadings	
	1	2
Embedded Figures Test	.81	
Un/segmented Block Design	.73	
Impossible Figures		.89
Navon Similarity Judgment		.62

The first factor contained the EFT and the Un/segmented Block Design Task. As both measures benefit from piecemeal processing, this factor was interpreted as *visual segmentation* (reversed). This is consistent with previous findings that have shown performance on EFT and block design tasks correlate highly in both TD and clinical groups, even when the effects of age and ability have been partialled out (Burnette et al., 2005; Jarrold et al., 2000; Ropar & Mitchell, 2001). The second factor included the preference towards matching by global form on the Navon Similarity Judgment Task and accurate performance on the Impossible Figures Task, and was interpreted as *visual integration.*

Composite scores to reflect each factor were constructed by averaging the respective standardised residual scores for each participant. The visual segmentation composite was significantly inversely related to the visual integration composite (r = -.21, p = .002). Interestingly, this relationship differed by gender; a significant negative relationship was found for female participants (r = -.30, p = .002), but no reliable correlation between the two composite scores emerged for male participants (r = -.06, p = .57; z_{r1-r2} = 1.76, p = .04). This finding implies that good integration ability went with good segmentation ability in females only, while the two abilities were independent in males.

ASD and Control Results

The inter-correlations between visual-perceptual domain variables are presented in Table 9.4 and Table 9.5 for the ASD and control groups, respectively. Based on the two-factor solution found in the TD sample, two composite scores were constructed to reflect the *visual segmentation* (reversed) and *visual integration* factors. As in the TD group, a significant negative correlation between the composite scores for visual segmentation and visual integration was evident in the control group (r = -.36, p = .05). However, in the ASD group a positive, but non-significant, correlation was found between the two composite scores (r = .12, p = .53; z_{r1-r2} = 1.86, p = .03), suggesting independence or even trade-off between visual segmentation and visual integration. That is, ASD participants good at segmentation were poor at integration; which was not found in TD or matched controls.

Table 9.4 Correlation matrix for low-level visuo-spatial coherence measures in the ASD group.

		Un/segmented Block Design	Impossible Figures	Navon Similarity Judgment
Embedded Figures Test	***r***	**.05**	**.01**	**.20**
	pr_{age}	.06	.01	.20
	$pr_{FIQ\text{-}BD}$	.05	.01	.21
	$pr_{age\ \&\ FIQ\text{-}BD}$	.06	.01	.21
Un/segmented Block Design	***r***		**-.10**	**.07**
	pr_{age}		-.08	.04
	$pr_{FIQ\text{-}BD}$		.02	.12
	$pr_{age\ \&\ FIQ\text{-}BD}$		.02	.09
Impossible Figures	***r***			**.18**
	pr_{age}			.19
	$pr_{FIQ\text{-}BD}$			.10
	$pr_{age\ \&\ FIQ\text{-}BD}$			.10

Table 9.5 Correlation matrix for low-level visuo-spatial coherence measures in the control group.

		Un/segmented Block Design	Impossible Figures	Navon Similarity Judgment
Embedded Figures Test	***r***	**.31~**	**-.49****	**-.18**
	pr_{age}	.33~	-.58**	-.18
	$pr_{FIQ\text{-}BD}$	.23	-.30	-.14
	$pr_{age\ \&\ FIQ\text{-}BD}$	.27	-.41*	-.13
Un/segmented Block Design	***r***		**-.42***	**.08**
	pr_{age}		-.39*	.07
	$pr_{FIQ\text{-}BD}$		-.35~	.11
	$pr_{age\ \&\ FIQ\text{-}BD}$		-.33~	.10
Impossible Figures	***r***			**.14**
	pr_{age}			.17
	$pr_{FIQ\text{-}BD}$			.08
	$pr_{age\ \&\ FIQ\text{-}BD}$			.11

~ p <.10, *p < .05 **p < .01

9.3 Low-level auditory domain

Typical Development Results

The inter-correlations among the low-level auditory coherence measures in the TD sample are presented in Table 9.6. In addition to age and IQ, number of years of music training was regressed out of the standardised residual score for three tasks that required music processing (see Table 9.6), prior to PCA. The purpose of this additional regression was to ensure that any correlations found between the measures were not simply due to general differences in music experience.

Table 9.6 Correlation matrix for the low-level auditory coherence measures in the TD group.

		Chord Segmentation	Pitch Identification	Chord Sequence
Phoneme Segmentation	r	**.23****	**.25****	**.00**
	pr_{age}	.22**	.24**	.01
	$pr_{FIQ\text{-}BD}$	.19**	.15*	.04
	pr_{music}	.19**	.21**	.04
	$pr_{age,\ FIQ\text{-}BD,\ music}$	.16*	.12~	.06
Chord Segmentation	r		**.32****	**-.09**
	pr_{age}		.30**	-.07
	$pr_{FIQ\text{-}BD}$		.27**	-.07
	pr_{music}		.21**	-.02
	$pr_{age,\ FIQ\text{-}BD,\ music}$		.17*	.00
Pitch Identification	r			**-.11**
	pr_{age}			-.08
	$pr_{FIQ\text{-}BD}$			-.08
	pr_{music}			-.03
	$pr_{age,\ FIQ\text{-}BD,\ music}$			.00

~ p <.10, *p < .05, **p < .01

In determining the factorability of the inter-correlation matrix, it was observed that Bartlett's test of sphericity did not reach significance (p = .07). Furthermore, the KMO index was .56 indicating that the degree of common variance between the four auditory measures was low. Taking an tentative approach, a PCA was conducted using varimax orthogonal rotation[10]. Two factors emerged with eigenvalues greater than 1.00, which together explained 59% of the variance (see Table 9.7). As the Chord Sequence Task did not correlate with the other auditory measures, the analysis was repeated excluding this

10 Nonorthogonal rotation methods were also used (e.g., direct oblimin) but yielded the same result suggesting that the underlying factors are orthogonal.

variable. Although the KMO index remained low (.57), Bartlett's test of sphericity was now significant (p = .002). Furthermore, the PCA conducted on the three remaining auditory measures produced one factor, which explained 44% of the variance. As only one component was extracted the solution could not be rotated. The resultant factor loadings for the second PCA are presented (in brackets) in Table 9.7.

Table 9.7 Factor loadings (> .40) from the PCA (orthogonal rotation) of the four (three) auditory-perceptual coherence measures in the TD sample with age and FIQ-BD removed.

Low-level auditory tasks	Factor Loadings	
	1	2
Chord Segmentation	.71 (.71)	
Pitch Identification	.67 (.66)	
Phoneme Segmentation	.60 (.62)	
Chord Sequence		.94

Note: Factor loadings (non-rotated) after excluding Chord Sequence presented in brackets.

In sum, the analyses suggested Chord Sequence fell on its own factor, while Phoneme Segmentation, Chord Segmentation and Pitch Identification all showed some degree of covariance through correlational and principal components analyses. The commonality between all three auditory measures is the ability to process the absolute properties of a sound (note or phoneme) to isolate and identify the stimulus independent of its presented context. The extracted factor was therefore interpreted as *auditory segmentation* (reversed).

No reliable correlation was found between the auditory segmentation composite (average of the standardised residual scores for Phoneme Segmentation, Chord Segmentation and Pitch Identification) and Chord Sequence coherence index in the TD sample overall (r = -.02, p = .85), or split by gender (males r = -.04, p = .77; females r = .02, p = .88). The coherence index from this task was the tendency to judge a final chord as sounding correct when it was congruent with the overall structure of a musical sequence yet incongruous with the last chord in the sequence. As well as an awareness of musical structure, decision-making processes are required to complete the task. Of note, Heaton (2006) refers to the Chord Sequence Task as a high-level measure of music cognition rather than a low-level perceptual task. As it could not be determined conclusively whether this task was tapping low-level auditory coherence processing, the coherence index for the Chord Sequence Task was taken in isolation to the next stage of between-domain analyses, alongside the low-level auditory composite.

ASD and Control Results

Table 9.8 and Table 9.9 present the inter-correlations among the four auditory coherence measures for the ASD and control groups respectively. Although the zero-order correlations between the Chord Segmentation, Phoneme Segmentation and Pitch Identification measures did not reach significance in the ASD or control groups, they were of similar magnitude to the TD group ($r = .10$ to $.39$). Interestingly, the strength of these correlations increased when the effects of age were removed in the control group ($pr = .17$ to $.33$), but were reduced when the effects of FIQ-BD were partialled out (all $pr < .08$). In contrast, partialling out age and FIQ-BD did not greatly impact the correlations in the ASD group.

The coherence index from the Chord Sequence Task showed a significant negative correlation with the Pitch Identification Task in the control group when age was taken into account ($pr = -.50$, $p = .02$), but not when the effects of IQ were removed ($pr = -.22$, $p = .34$). The direction of this correlation suggested participants who showed a tendency to process chord sequences globally also showed good ability to identify isolated notes, a skill thought to require local processing. The correlation between the two measures did not reach significance in the ASD group ($r = -.11$ to $-.08, p > .62$; $z_{r1-r2} = 1.00, p = .16$).

An association was observed in the ASD group between the Chord Sequence and Chord Segmentation Tasks ($r = .36$ to $.39$, all $p < .10$). Although not statistically significant, the strength of correlation coefficients were classified as moderate (J. Cohen, 1988), and suggested that individuals with weaker global bias in processing chord sequences showed greater accuracy in segmenting a single tone from a chord. Thus a trade-off between local and global processing was somewhat indicated in the ASD group.

In accordance with the data reduction approach taken in the TD group, auditory segmentation composite scores were calculated for each participant by averaging the standardised residual scores for the Chord Segmentation, Phoneme Segmentation and Pitch Identification Tasks, which took into account the effects of age and ability (and years of music training). As found in the TD group, no reliable correlation was observed between the auditory segmentation composite score and the Chord Sequence coherence index in the ASD ($r = -.11, p = .63$) or control group ($r = -.12, p = .61$).

Table 9.8 Correlation matrix for low-level auditory coherence measures in the ASD group.

		Chord Segmentation	Pitch Identification	Chord Sequence
Phoneme Segmentation	r	**.10**	**.18**	**-.25**
	pr_{age}	.12	.20	-.35
	pr_{FIQ-BD}	.15	.28	-.39~
	pr_{music}	.17	.17	.26
	$pr_{age, FIQ-BD, music}$	.25	.29	-.47*
Chord Segmentation	r		**.31~**	**.36**
	pr_{age}		.31~	.37~
	pr_{FIQ-BD}		.31	.38~
	pr_{music}		.33~	.36
	$pr_{age, FIQ-BD, music}$		.33~	.39~
Pitch Identification	R			**-.11**
	pr_{age}			-.11
	pr_{FIQ-BD}			-.09
	pr_{music}			-.11
	$pr_{age, FIQ-BD, music}$			-.08

~ p <.10, *p < .05

Table 9.9 Correlation matrix for low-level auditory coherence measures in the control group.

		Chord Segmentation	Pitch Identification	Chord Sequence
Phoneme Segmentation	r	**.16**	**.39***	**.04**
	pr_{age}	.17	.33~	-.05
	pr_{FIQ-BD}	-.10	.03	.31
	pr_{music}	.01	.28	.15
	$pr_{age, FIQ-BD, music}$	-.20	-.12	.34
Chord Segmentation	r		**.29**	**-.22**
	pr_{age}		.32~	-.23
	pr_{FIQ-BD}		.08	-.12
	pr_{music}		.18	-.16
	$pr_{age, FIQ-BD, music}$		.02	-.03
Pitch Identification	r			**-.36~**
	pr_{age}			-.50*
	pr_{FIQ-BD}			-.22
	pr_{music}			-.31
	$pr_{age, FIQ-BD, music}$			-.32

~ p <.10, *p < .05

9.4 High-level visuo-spatial domain

Typical Development Results

Table 9.10 presents the inter-correlations among the coherence measures from the high-level visuo-spatial domain in the TD group. As seen in the table, very few significant correlations between variables were found. Of note, two measures of global processing (coherent drawing style and global descriptions of the underwater scene) were found to have a significant positive correlation, although the correlation did not survive partialling out age and FIQ-BD.

Individuals who tended to draw in a globally coherent manner, also tended to show good surface form retention on the Picture Memory Task, a measure hypothesised to reflect effective local processing. This correlation remained significant after the effects of age and IQ were controlled for, and suggests that TD participants have the capacity to switch between the local/global processing demands.

Individuals who showed local coherence on the Fragmented Pictures Task (i.e., more of the picture was required before recognition of a degraded picture could be made) were more likely to provide global descriptions of the underwater scene. This inverse relationship between two measures hypothesised to access global processing is not easily explained, but it is notable that the relationship only emerges when age is partialled out.

As the high-level visuo-spatial measures were deemed unsuitable for formal data reduction methods (KMO = .49, Bartlett's test of sphericity p = .007), all coherence indices were carried over to the between-domain analyses. The exception to this was the Picture Memory Task where the average of the two coherence indices originating from this task was used (global description and discrimination of surface form changes) to avoid testing multiple indices from a single task.

Table 9.10 Correlation matrix for high-level visuo-spatial coherence measures in the TD group

		Picture Memory: Description	Picture Memory: Recognition	Drawing style
Fragmented Pictures	***r***	**.01**	**-.09**	**.10**
	pr_{age}	-.17*	-.01	-.09
	$pr_{FIQ\text{-}BD}$	.01	-.09	.10
	$pr_{age\ \&\ FIQ\text{-}BD}$	-.18*	.00	-.09
Picture Memory: Description	***r***		**-.10**	**.14***
	pr_{age}		-.04	.02
	$pr_{FIQ\text{-}BD}$		-.07	.13~
	$pr_{age\ \&\ FIQ\text{-}BD}$		-.01	.00
Picture Memory: Recognition	***r***			**-.26****
	pr_{age}			-.22**
	$pr_{FIQ\text{-}BD}$			-.26**
	$pr_{age\ \&\ FIQ\text{-}BD}$			-.21**

~ p <.10, *p < .05, **p < .01

ASD and Control Results

Table 9.11 and Table 9.12 present the inter-correlations among the high-level visuo-spatial coherence indices for the ASD and control groups respectively. Although few significant correlations are observed overall, a relationship was found in the ASD group between providing global descriptions of the underwater scene and quicker identifications of objects on the Fragmented Pictures Task. Thus, the ability to integrate information to form a coherent whole was evident across two visuo-spatial tasks in the ASD group. This relationship held significance after the effects of age (but not FIQ-BD) were partialled out. Interestingly, the inverse was observed in the control group (and TD sample) whereby a significant negative correlation was found between the two measures, suggesting individuals who showed good visual integration on the Fragmented Pictures Task tended not to describe the overall theme of a complex scene.

It appears that coherence may be driving performance in the ASD, but other processes/strategies could be operating in control and TD individuals. For example, the scene description placed demands on verbal abilities, in addition to visual-integration. Longer detection times on the Fragmented Pictures Task may result from a more cautious approach rather than poor visual-integration. Furthermore, the on-line processing demands of the Fragmented Pictures Task (to cohere degraded information) are in contrast to the off-line processing requirements of the Picture Memory Task (to describe a previously viewed picture).

Table 9.11 Correlation matrix for high-level visuo-spatial coherence measures in the ASD group

		Picture Memory: Description	Picture Memory: Recognition	Drawing style
Fragmented Pictures	***r***	**.38***	**.08**	**.00**
	pr_{age}	.38*	.07	.01
	$pr_{FIQ\text{-}BD}$	.20	.21	-.08
	$pr_{age\ \&\ FIQ\text{-}BD}$	.21	.20	-.07
Picture Memory: Description	***r***		**-.07**	**.07**
	pr_{age}		-.05	.06
	$pr_{FIQ\text{-}BD}$		.04	.00
	$pr_{age\ \&\ FIQ\text{-}BD}$		.05	-.01
Picture Memory: Recognition	***r***			**-.11**
	pr_{age}			-.07
	$pr_{FIQ\text{-}BD}$			-.07
	$pr_{age\ \&\ FIQ\text{-}BD}$			-.05

$^{*}p < .05$

Table 9.12 Correlation matrix for high-level visuo-spatial coherence measures in the control group

		Picture Memory: Description	Picture Memory: Recognition	Drawing style
Fragmented Pictures	***r***	**-.28**	**-.08**	**.19**
	pr_{age}	-.28	-.09	.20
	$pr_{FIQ\text{-}BD}$	-.36~	-.05	.17
	$pr_{age\ \&\ FIQ\text{-}BD}$	-.40*	-.03	.17
Picture Memory: Description	***r***		**-.27**	**.05**
	pr_{age}		-.27	.05
	$pr_{FIQ\text{-}BD}$		-.16	-.03
	$pr_{age\ \&\ FIQ\text{-}BD}$		-.15	-.03
Picture Memory: Recognition	***r***			**-.26**
	pr_{age}			-.26
	$pr_{FIQ\text{-}BD}$			-.22
	$pr_{age\ \&\ FIQ\text{-}BD}$			-.21

$\sim p < .10$, $^{*}p < .05$

As in the TD group, control participants who tended to draw coherently also showed good surface form retention on the Picture Memory Task, although the correlation coefficients did not reach significance (r = -.21 to -.26). No association was found between drawing style and surface form retention in the ASD group (r = -.05 to -.11).

As no clear task clustering emerged from the TD data, the coherence indices from each visuo-spatial task (with the average of the two Picture Memory indices) were taken to the next stage of between-domain analyses.

9.5 High-level verbal domain

Typical Development Results

The inter-correlations among the coherence indices from the high-level verbal domain tasks in the TD sample are presented in Table 9.13. As no significant correlations were observed, the data were considered unsuitable for principle components analysis. Indeed the verbal coherence indices showed unsatisfactory values regarding the KMO index (.48, classified as unacceptable according to Kaiser, 1974), and failed Bartlett's test of sphericity (p = .38). As the verbal measures purporting to tap central coherence processing were unsuitable for formal data reduction methods, all coherence indices were carried through to the between-domain analyses, with the exception of the Story Memory Task which used the average of the two coherence indices (proportion of verbatim recall and discrimination of surface form changes).

Table 9.13 Correlation matrix for high-level verbal coherence measures in the TD group

		Homographs: Difference Score	Story Memory: Recognition	Story Memory: Verbatim Recall
Sentence Completion	r	**-.11**	**-.05**	**-.02**
	pr_{age}	-.07	-.04	-.05
	$pr_{FIQ\text{-}BD}$	-.11	-.05	-.02
	$pr_{age\ \&\ FIQ\text{-}BD}$	-.07	-.04	-.05
Homographs: Difference Score	r		**-.10**	**-.03**
	pr_{age}		-.12	-.01
	$pr_{FIQ\text{-}BD}$		-.11	-.04
	$pr_{age\ \&\ FIQ\text{-}BD}$		-.13~	-.02
Story Memory: Recognition	r			**.02**
	pr_{age}			.03
	$pr_{FIQ\text{-}BD}$			.00
	$pr_{age\ \&\ FIQ\text{-}BD}$			.01

~ p <.10

ASD and Control Results

The inter-correlations between verbal-semantic domain variables are presented in Table 9.14 and Table 9.15 for the ASD and control groups, respectively. A significant negative correlation was observed in the control group between the Sentence Completion Task and the Story Memory Task for surface form recognition that held after the effects of age and ability were partialled out. This inverse relationship suggested that control group participants who tended to provide global completions to sentence stems also showed good ability to discriminate sentences that differed in surface form. While this potentially shows an ability to switch between local/global task demands, an underlying general ability factor may also be driving results, which was undetected by the IQ measure used. This correlation was not observed in the ASD group; instead a positive, although non-significant, correlation was found between the two measures (r = .16 to .21).

Table 9.14 Correlation matrix for high-level verbal coherence measures in the ASD group

		Homographs: Difference Score	Story Memory: Recognition	Story Memory: Verbatim Recall
Sentence Completion	***r***	**.11**	**.16**	**-.22**
	pr_{age}	.10	.16	-.22
	$pr_{FIQ\text{-}BD}$	.29	.21	-.28
	$pr_{age\ \&\ FIQ\text{-}BD}$	.30	.21	-.28
Homographs: Difference Score	***r***		**.03**	**-.21**
	pr_{age}		.03	-.20
	$pr_{FIQ\text{-}BD}$		.01	-.18
	$pr_{age\ \&\ FIQ\text{-}BD}$		.00	-.18
Story Memory: Recognition	***r***			**.05**
	pr_{age}			.05
	$pr_{FIQ\text{-}BD}$			.05
	$pr_{age\ \&\ FIQ\text{-}BD}$			.06

Table 9.15 Correlation matrix for high-level verbal coherence measures in the control group

		Homographs: Difference Score	Story Memory: Recognition	Story Memory: Verbatim Recall
Sentence Completion	***r***	**-.17**	**-.42***	**.31**
	pr_{age}	-.16	-.63**	.25
	$pr_{FIQ\text{-}BD}$	-.17	-.45*	.34~
	$pr_{age\ \&\ FIQ\text{-}BD}$	-.12	-.63**	.31
Homographs: Difference Score	***r***		**.20**	**-.11**
	pr_{age}		.25	-.10
	$pr_{FIQ\text{-}BD}$		.05	-.31
	$pr_{age\ \&\ FIQ\text{-}BD}$		.14	-.28
Story Memory: Recognition	***r***			**-.06**
	pr_{age}			-.22
	$pr_{FIQ\text{-}BD}$			-.23
	$pr_{age\ \&\ FIQ\text{-}BD}$			-.33~

~ p <.10, *p < .05, **p < .01

Although not reaching statistical significance, moderate positive correlations ($r = .25$ to $.34$) were found between Sentence Completion and Story Memory (verbatim recall) in the control group. This suggested that individuals who tended to complete sentence stems with a local associate, also tended to recall stories verbatim. The relationship between variables was opposite in the ASD group and, although not reaching significance ($r = -.22$ to $-.28$), the correlations were of different magnitude to the control group ($z_{r1-r2} = 2.04$, $p = .02$). The expected tendency for ASD participants to recall stories verbatim was not found in the experimental data, which may account for the lack of unity across verbal tasks in the clinical group.

As no task clustering was apparent from the TD data, the coherence indices from each verbal coherence task (with the average of the two Story Memory indices) were taken forward to the between-domain analyses.

9.6 Between-domain analyses

Table 9.16 presents the inter-correlations between the key coherence indices selected from each processing-level/modality quadrant for the TD sample (the correlations by gender can be found in Appendix P). The same correlation matrix is presented for the ASD and control groups in Table 9.17. The following provides a summary of the pattern of correlations found across low-level and high-level tasks, and across visuo-spatial and auditory-verbal modalities. Consistency of coherence processing style across both level and modality is also examined. Findings from the PCA conducted on all coherence indices are also presented.

9.6.1 Low-level tasks across modality

There was some evidence of cross-modality processing style at low-levels: visual segmentation composite correlated with auditory segmentation composite in the TD sample ($r = .14$, $p = .05$). This relationship was significant in female participants ($r = .22$, $p = .03$), but not for male participants ($r = .07$, $p = .51$; difference between correlations not significant, $z_{r1-r2} = 1.07$, $p = .28$). While visual segmentation and auditory segmentation composites correlated significantly in the control group ($r = .40$, $p = .03$), there was no such relationship in the ASD group ($r = -.08$, $p = .68$; difference between coefficients marginally significant, $z_{r1-r2} = 1.89$, $p = .06$).

No association was found between the Chord Sequence Task and the visual integration composite in any group which questions whether the Chord Sequence Task is a valid measure of global processing at a low-level. Furthermore, significant negative correlations were found between the Chord Sequence Task and the visual segmentation composite in the TD sample (all $r = -.23$, $p = .005$; males $r = -.24$, $p = .04$; females $r = -.23$,

p = .04), suggesting individuals who perceived the global structure of chord sequences were also well able to segment visual information. The correlation coefficients were of the same magnitude in the ASD (r = -.20, p = .37) and control groups (r = -.25, p = .27), although not statistically significant.

Significant negative correlations were found in the ASD (r = -.41, p = .03) and control groups (r = -.35, p = .05) between visual integration and auditory segmentation composites, which approached significance in the TD group (r = -.13, p = .07). Those showing an ability to integrate visuo-spatial information were able to proficiently segment auditory information, thereby showing an ability to be both a good local processor (of visual information) and a good global processor (of auditory stimuli).

9.6.2 *High-level tasks across modality*

The evidence for cross-modality processing style at high-levels was scant in both the TD sample and ASD and control groups. One association to note was a negative correlation between Sentence Completion and Picture Memory (r = -.16, p = .03; trend in males r = -.19, p = .07; not significant in females r = -.11, p = .26; z_{r1-r2} = 0.58, p = .57). Participants who tended to provide global completions to sentence stems showed local processing in the Picture Memory Task (i.e., better recognition memory and not global descriptions), a relationship not in the direction predicted by central coherence theory. This relationship was found in the control group (r = -.40, p = .03), but not in the ASD group where a positive (non-significant) correlation was found (r = .18, p = .35). There was also a statistically significant difference in the strength of the correlations for the ASD and control groups (z_{r1-r2} = 2.27, p = .02).

9.6.3 *Visuo-spatial modality across levels*

Lower-level composites correlated in the predicted direction with some higher-level tasks in the visuo-spatial modality. Visual integration correlated with Fragmented Pictures in the TD group (r = .14, p = .05), particularly in males (r = .24, p = .02), but not females (r = .05, p = .62; z_{r1-r2} = 1.37, p = .17). Individuals who showed good visual integration on low-level tasks, also showed good visual integration with more meaningful material. This association was also found in the control group (r = .22, p = .23), although not at a significant level, and to a lesser degree in the ASD group (r = .12, p = .54; z_{r1-r2} = 0.38, p = .70).

Although not found in the TD sample, a significant positive correlation was found between visual segmentation composite and Picture Memory in both the ASD (r = .39, p = .03) and control groups (r = .38, p = .03). This may show consistency in local processing across both indices: individuals who showed high visual segmentation skills also showed an ability to discriminate surface form changes, and tended not to provide global descriptions.

Correlations opposite to predictions from central coherence theory also emerged in the control group: Coherent drawing style was related to better visual segmentation abilities (r = -.42, p = .02), and good visual integration was related to local processing on the Picture Memory Task (r = -.41, p = .02). Control participants thus showed an ability to be good at both local processing and global processing.

9.6.4 Auditory/verbal modality across levels

Some evidence of consistency of cognitive style across processing level was found in auditory/verbal modality for TD participants: Homograph Reading correlated with auditory segmentation composite (r = .20, p = .004). Thus, individuals who were sensitive to context to disambiguate homographs were less able to process the absolute properties of a sound. This relationship was significant for males (r = .25, p = .02) but not for females (r = .15, p = .12; $z_{r1\text{-}r2}$ = 0.73, p = .46). Furthermore, the correlation approached significance in the control group (r = .36, p = .06), but was somewhat reduced in the ASD group (r = .16, p = .39; $z_{r1\text{-}r2}$ = 0.80, p = .21).

Contrary to predictions of central coherence, the tendency to process the global properties of music was associated with local processing on the Story Memory Task in control participants (r = -.42, p = .05). No such relationship was observed in the ASD group (r = -.07, p = .77; $z_{r1\text{-}r2}$ = 1.16, p = .12) or the TD sample (r = -.11, p = .19).

9.6.5 Cross-modality and cross-level

There was insufficient evidence to show the unity of individual differences in coherence across both level and modality. No clear picture emerged from the pattern of correlations found in the TD sample, nor in the ASD and control groups. One significant negative correlation suggested that TD participants with good visual integration abilities were more likely to show local processing on the Story Memory Task (r = -.16, p = .03). This relationship was significant in females (r = -.23, p = .02), but not in males (r = -.10, p = .31; $z_{r1\text{-}r2}$ = 0.94, p = .35). Additionally, a negative trend in the ASD group suggested that global processing on the Fragmented Pictures Task did not necessarily relate to global processing of musical sequences (r = -.42, p = .052). The reverse was shown in the control group by a positive, but non-significant, correlation between the two measures (r = .24, p = .29; $z_{r1\text{-}r2}$ = 2.13, p = .02), while no association was found in the TD sample (r = .02, p = .82).

Table 9.16 Between-domain Pearson correlations of coherence measures (standardised residuals with effects of age and FIQ-BD removed) for the TD group.

		Low-level auditory	Low-level visuo-spatial		High-level visuo-spatial			High-level verbal		
		Chord Sequence	Segment	Integrate	Fragment	Picture Memory	Drawing Style	Sentence	Homo-graphs[a]	Story Memory
Low-level auditory	Segmentation	.02	.14*	-.13~	-.03	-.02	-.05	-.04	.20**	.00
	Chord Sequence		-.23**	-.01	.02	-.02	-.03	.09	-.14~	-.11
Low-level visuo-spatial	Segmentation			-.21**	-.06	.01	-.02	-.04	.06	.08
	Integration				.14*	-.07	-.03	.09	.08	-.16*
High-level visuo-spatial	Fragment Pictures					-.13~	-.09	-.05	.08	-.05
	Picture Memory						-.16*	-.16*	-.08	.01
	Drawing Style							.04	.01	-.07
High-level verbal	Sentence Completion								-.10	-.06
	Homographs[a]									-.12

~ p <.10, *p < .05 **p < .01
[a]Extreme outliers removed

Table 9.17 Between-domain Pearson correlations of coherence measures (standardised residuals with effects of age and FIQ-BD removed) for the ASD and control groups.

		Low-level auditory	Low-level visuo-spatial		High-level visuo-spatial			High-level verbal		
ASD		Chord Sequence	Segment	Integrate	Fragment	Picture Memory	Drawing Style	Sentence	Homographs	Story Memory
Low-level auditory	Segmentation	-.11	-.08	-.41*	-.31~	-.13	-.10	.18	.16	.31~
	Chord Sequence		-.20	-.05	-.42~	-.13	.22	-.18	-.36	-.07
Low-level visuo-spatial	Segmentation			.12	.15	.39*	.02	.08	-.09	.07
	Integration				.12	-.19	.13	-.02	.28	-.11
High-level visuo-spatial	Fragmented Pictures					.28	-.06	.03	.24	.06
	Picture Memory						-.02	.18	-.11	.06
	Drawing Style							-.08	-.22	-.20
High-level verbal	Sentence Completion								.25	-.05
	Homographs									-.12

Control		Chord Sequence	Segment	Integrate	Fragment	Picture Memory	Drawing Style	Sentence	Homographs	Story Memory
Low-level auditory	Segmentation	-.11	.40*	-.35*	-.04	.04	-.07	.02	.36~	.00
	Chord Sequence		-.25	.06	.24	-.01	-.01	.03	.19	-.42*
Low-level visuo-spatial	Segmentation			-.36*	-.14	.38*	-.42*	-.29	.16	.19
	Integration				.22	-.41*	-.06	.15	-.34~	-.26
High-level visuo-spatial	Fragmented Pictures					-.33~	.17	-.09	.16	-.03
	Picture Memory						-.19	-.40*	.12	.04
	Drawing Style							.32~	.25	-.18
High-level verbal	Sentence Completion								-.10	-.27
	Homographs									-.07

~ $p < .10$, * $p < .05$

9.6.6 *Principal component analyses*

Despite the inconsistencies in the pattern of correlations within and between each processing-level/domain, the coherence indices from the TD sample were subjected to a PCA to determine any underlying factor structure. The Drawing Style index was removed from analysis due to low inter-task correlations and low individual KMO index (.42). The factorability of the remaining nine coherence indices was confirmed by Bartlett's test of sphericity (p = .05). However, the overall KMO index was .52 suggesting that the degree of common variance was low. With this in mind, a PCA was conducted using varimax orthogonal rotation. Four factors were found, which together explained 57% of the variance in the data (see Table 9.18).

Table 9.18 Factor loadings (> .40) from the PCA (orthogonal rotation) using nine coherence measures in the TD sample with age and FIQ-BD removed.

	Factor Loadings			
	1	2	3	4
Visual integration composite	.70			
Fragmented Pictures	.59			
Chord Sequence		.76		
Visual segmentation composite		-.60		
Story Memory		-.52		
Auditory segmentation composite			.73	
Homographs			.72	
Picture Memory				-.75
Sentence Completion				.74

The first factor, incorporating both low-level visual integration composite and Fragmented Pictures, was interpreted as *general* visual-integration. This common factor may indicate that the Fragmented Pictures Task tapped visual-integration abilities at a lower level than predicted, or that there is some consistency in coherence across high and low processing levels.

In the second factor, Story Memory was found to cluster with the visual segmentation composite in a positive direction. A commonality across these measures is the demand for local processing for 'good' performance (i.e., discrimination of surface form changes in sentences, tendency for verbatim recall, less benefit of segmentation in the construction of block designs, and fast/accurate disembedding on the EFT). The second factor also includes a *negative* relationship to Chord Sequence, which was predicted to assess the degree of global processing of musical sequences. The asymmetry shown in the second factor may therefore indicate the propensity for TD individuals to demonstrate both good global processing and segmentation/featural processing.

As found in the correlations within the auditory/verbal modality, the third factor suggests an association between auditory segmentation ability and reduced effect of context position in the reading of homographs. This factor may be interpreted as poor auditory/verbal integration.

The fourth factor that emerged appeared to tap verbal abilities in describing a complex scene and finding a completion to a sentence stem. Visual recognition memory for local details was also assessed, although this can be verbally mediated. The opposing direction of the two indices in the factor solution suggested that the extent of coherent completions made to sentence stems (global processing) was related to the ability to recognise surface changes and the tendency to not provide a global description of the complex scene (local processing). Although the effects of age and IQ were removed from analysis, an underlying ability or strategic factor may still drive performance on these measures and explain the absence of trade-off between segmentation and integration processes in the TD group.

As the PCA included three composite variables (visual integration, visual segmentation, auditory segmentation), a second PCA was run that included only the original coherence indices in order to check the validity of collapsing the data. The factorability of the data was confirmed by Bartlett's test of sphericity (p = .01), although the KMO index was .53, indicating the common variance between variables was low. Six factors emerged from the PCA, which together explained 61% of the variance (see Table 9.19).

Common to the original PCA, was the clustering of the three coherence indices that comprise the auditory segmentation composite (Phoneme Segmentation, Pitch Identification, Chord Segmentation) together with the effect of context position in the reading of homographs, resulting in an auditory/verbal integration factor (Factor 2). Furthermore, the negative relationship between Sentence Completion and Picture Memory remained (Factor 3).

Table 9.19 Factor loadings (> .40) from the PCA (orthogonal rotation) using thirteen coherence measures in the TD sample with age and FIQ-BD removed.

	Factor Loadings					
	1	2	3	4	5	6
Chord Sequence	-.81					
Embedded Figures Test	.60					
Phoneme Segmentation		.68				
Pitch Identification		.64				
Chord Segmentation		.53				
Homographs		.41				
Sentence Completion			.74			
Picture Memory			-.63			
Fragmented Pictures				-.83		
Navon Similarity Judgment					.74	
Story Memory					-.59	
Un/segmented Block Design						.79
Im/possible Figures						-.71

However, the coherence indices that comprise the visual integration and visual segmentation composites did not fall into common factors. Instead, the asymmetry of the direction of factor loadings indicates the tendency to show both good local and good global processing. Thus, the tendency to judge global completions as sounding correct on the Chord Sequence Task was associated with fast identifications on the EFT (Factor 1), global matching on the Navon Similarity Judgment Task was associated with detail processing on the Story Memory Task (Factor 5), and a lack of benefit of segmentation on the Un/segmentation Block Design Task was related to good ability to determine geometric impossibility (Factor 6). Finally, the Fragmented Pictures Task emerged as an isolated factor (Factor 4).

9.7 Subgroup Analysis

The predominant finding from the pattern of correlations and PCA (primarily in the TD sample) is the suggestion that good local processing is not necessarily achieved at the cost of poor global processing. As coherence tasks were deliberately selected/designed to assess independently the demands of both local and global processing, this suggests that individuals were able to adapt processing style and respond locally or globally according to which the task demanded. Individuals may vary in degree of this adaptability, and also show predominance in either local or global processing (bias) irrespective of task demands.

To study this proposal further, a subgroup analysis was undertaken. Firstly, tasks were classified as *global adaptive* or *local adaptive* according to whether processing demands were primarily on global or local strategies/abilities. For example, Pitch Identification, Phoneme/Chord Segmentation, EFT, Un/segmented Block Design, and recognition memory on Story/Picture Memory Tasks, all place demands on (can benefit from) local processing. Fragmented Pictures, Homograph Reading, and Impossible Figures require proficient global processing. Several tasks were designed to be 'open-ended' such that taking a global or local approach did not relate directly to task success, or that a 'correct' answer did not exist per se (e.g., Sentence Completion, Drawing Style, Navon Similarity Judgment Task, Chord Sequence). However, developmental effects were typically found on such tasks, indicating a local to global progression in the TD sample, and thus these open-ended tasks were classified as global adaptive.

Each participant received a local score on global adaptive tasks, and a local score on local adaptive tasks, with each score reflecting the number of times s/he had been categorised as showing weak central coherence[11], averaged over the number of tasks completed (maximum 9 for global adaptive, and 8 for local adaptive). Participants were then categorised into one of four subgroups: Local Dominant if their mean local score was above the 67^{th} percentile of the TD sample on local adaptive tasks ($\geq .38$) and on global adaptive tasks ($\geq .33$); Global Dominant if their mean local score for both local adaptive and global adaptive tasks were below the 67^{th} percentile; Good Adaptive if their mean local score was above the 67^{th} percentile on local adaptive tasks and below the 67^{th} percentile on global adaptive tasks; Poor Adaptive if their mean local score was below the 67^{th} percentile on local adaptive tasks and above the 67^{th} percentile on global adaptive tasks (see Table 9.20). Participants were thus characterised on whether their task performance was predominantly a local or global style across all tasks (Local Dominant, Global Dominant), or whether they showed a degree of adaptability to the tasks demands (Good Adaptive, Poor Adaptive). Participant characteristics for TD, ASD and control groups categorised by this method are presented in the following sections.

[11] Based on the categorical summary tables presented at the end of Chapters 5 through to 8 (Table 5.19, Table 6.16, Table 7.18, Table 8.11).

Table 9.20 Four categories proposed to characterise individual performance across locally and globally adaptive coherence tasks.

		Global Adaptive Coherence Tasks	
		Local score $< .33$	*Local score* $\geq .33$
Local Adaptive Coherence Tasks	*Local score* $< .38$	Global Dominant	Poor Adaptive
	Local score $\geq .38$	Good Adaptive	Local Dominant

9.7.1 Typical Development subgroups

Table 9.21 presents the characteristics of all TD participants categorised into the four subgroups, and Figure 9.1 shows the percentage (and number) of male and female participants in each age group classified by subgroup. Developmental effects can be seen in degree of adaptability, with a clear shift in the proportion of participants classified as poor adaptors to good adaptors with age. This pattern is seen in both male and female TD participants. Independent samples t-tests found good adaptive participants were significantly older and had higher IQ than poor adaptive participants, in both male and female samples (all $t > 3.07$, $p < .004$).

In general, TD participants who showed predominately global processing across measures were older than those who showed predominately local processing ($t_{(53.7)} = 3.06$, $p = .003$, unequal variances), although the two subgroups were equivalent on all IQ measures (all $t_{(62)} < 1.36$, $p > .17$). The age effect was significant in female participants, with global responders being older ($M = 15.8$, $SD = 6.0$) than local responders ($M = 11.2$, $SD = 1.9$; $t_{(32)} = 2.80$, $p = .009$). The age difference did not reach significance in male participants (global responders $M = 15.2$, $SD = 3.5$, local responders, $M = 13.4$, $SD = 3.7$; $t_{(28)} = 1.38$, $p = .18$). While no difference in IQ was found between females who showed a predominately global or local response style (all $t_{(32)} < 1.06$, $p > .28$), IQ effects were found in male participants, such that males showing a local response style had higher FIQ ($t_{(28)} = 2.06$, $p = .05$) and FIQ-BD ($t_{(28)} = 2.31$, $p = .03$).

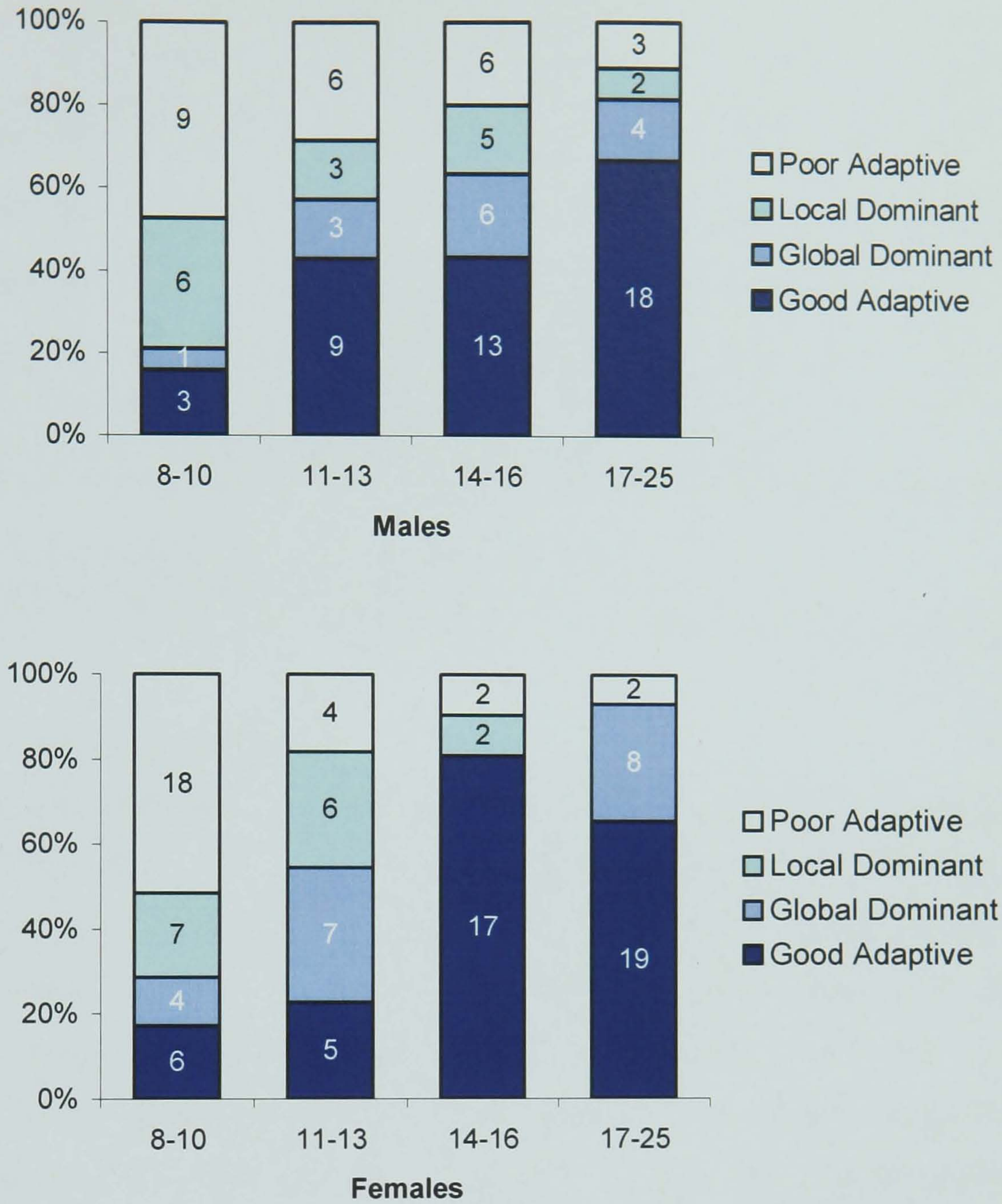


Figure 9.1 Bar charts showing the percentage (and number) of male and female participants in each age group categorised as global responders, local responders, good adaptive, and poor adaptive.

Table 9.21 TD sample characteristics by subgroup: Mean (*SD*)

Subgroup	*N*	Age[a]	FIQ[b]	VIQ[c]	PIQ[d]	FIQ-BD[e]
Poor Adaptive	50	12.2 (3.7)	97.7 (13.8)	101.7 (15.2)	94.0 (13.1)	100.0 (13.7)
Local Dominant	31	12.3 (3.1)	107.7 (15.1)	109.6 (18.9)	103.3 (11.1)	108.2 (16.2)
Global Dominant	33	15.5 (5.1)	104.5 (12.7)	104.9 (15.3)	102.5 (12.9)	103.1 (13.1)
Good Adaptive	90	16.5 (3.9)	113.9 (11.8)	115.3 (12.8)	108.6 (12.2)	112.8 (11.8)

[a]Poor Adaptive, Local Dominant < Global Dominant, Good Adaptive, $p < .01$
[b]Poor Adaptive < Local Dominant, Good Adaptive, $p < .006$; Global Dominant < Good Adaptive, $p = .002$
[c]Poor Adaptive, Global Dominant < Good Adaptive, $p < .005$
[d]Poor Adaptive < Local Dominant, Global Dominant, Good Adaptive, $p < .04$
[e]Poor adaptive < Local Dominant, Good Adaptive, $p < .04$; Global Dominant < Good Adaptive, $p = .002$.

9.7.2 ASD and Control subgroups

Participant characteristics by subgroup are presented in Table 9.22 for the ASD group and Table 9.23 for the control group. The percentage (and number) of ASD and control participants categorised by subgroup is presented in Figure 9.2. A higher proportion of ASD participants were classified as showing a predominantly local response style (23%, 7 from 31), compared to the matched controls (6%, 2 from 31; $\chi^2 = 3.25$, $p = .03$, one-tailed). No other subgroup varied significantly in the relative proportion of participants across ASD and control groups (all $\chi^2 < 1.19$, $p > .14$). Of the seven ASD participants classified as local responders, three had received a diagnosis of autism (and three were in the low IQ subgroup). Two of the 10 ASD participants classified as poorly adaptive, also had a diagnosis of autism. All ASD participants classified as global responders or highly adaptive had received a diagnosis of Asperger syndrome.

Within group analyses found participants did not differ in age across the four subgroups in the ASD group (Kruskal Wallis Test, $\chi^2 = 1.64$, $p = .65$), and control group ($\chi^2 = 1.22$, $p = .75$). Effects of IQ were significant in the control group (all $\chi^2 > 9.00$, $p < .03$), and indicated that those with high IQ showed good adaptability to respond locally or globally according to task demands. Overall effects of IQ only approached significance in the ASD group (all $\chi^2 < 7.58$, $p > .06$), possibly due to the high variance in IQ scores (with the exception of the predominately global responders). Pairwise comparisons did, however, reveal IQ differences between local and global responders in the ASD group, with local responders showing lower FIQ, VIQ and FIQ-BD (all $z > 2.13$, $p < .04$). This is opposite to males in the TD sample, where local responders had significantly higher IQ than global responders. Between group analyses showed no differences in age and IQ for ASD and control participants classified by subgroup (all $z < 1.47$, $p > .13$).

Table 9.22 ASD group characteristics by subgroup: Mean (*SD*)

Subgroup	*N*	Age	FIQ	VIQ	PIQ	FIQ-BD
Poor Adaptive	10	15.1 (3.6)	92.3 (22.3)	96.9 (22.6)	88.4 (19.5)	95.4 (19.9)
Local Dominant	7	15.3 (1.6)	77.4 (28.1)	81.5 (28.7)	78.4 (23.7)	75.1 (21.3)
Global Dominant	6	14.8 (1.6)	102.9 (5.8)	105.0 (2.2)	99.4 (10.8)	104.9 (6.0)
Good Adaptive	8	14.0 (1.5)	102.2 (24.2)	103.3 (23.2)	99.8 (23.4)	106.3 (24.8)

Table 9.23 Control group characteristics by subgroup: Mean (*SD*)

Subgroup	*N*	Age	FIQ[a]	VIQ[b]	PIQ[c]	FIQ-BD[d]
Poor Adaptive	12	15.2 (2.5)	78.7 (20.6)	85.2 (23.3)	77.8 (15.8)	81.8 (22.3)
Local Dominant	2	15.8 (0.7)	94.4 (12.4)	87.9 (5.7)	102.9 (16.4)	92.7 (13.4)
Global Dominant	5	15.0 (2.3)	93.5 (18.2)	96.8 (24.3)	91.7 (11.8)	92.9 (19.9)
Good Adaptive	12	14.2 (2.5)	109.7 (11.1)	110.8 (12.6)	105.5 (10.1)	108.7 (11.2)

[a]Poor Adaptive, Global Dominant < Good Adaptive, $p < .04$
[b]Poor Adaptive, Local Dominant < Good Adaptive, $p < .04$
[c]Poor Adaptive, Global Dominant < Good Adaptive, $p < .04$
[d]Poor Adaptive, Global Dominant < Good Adaptive, $p < .05$

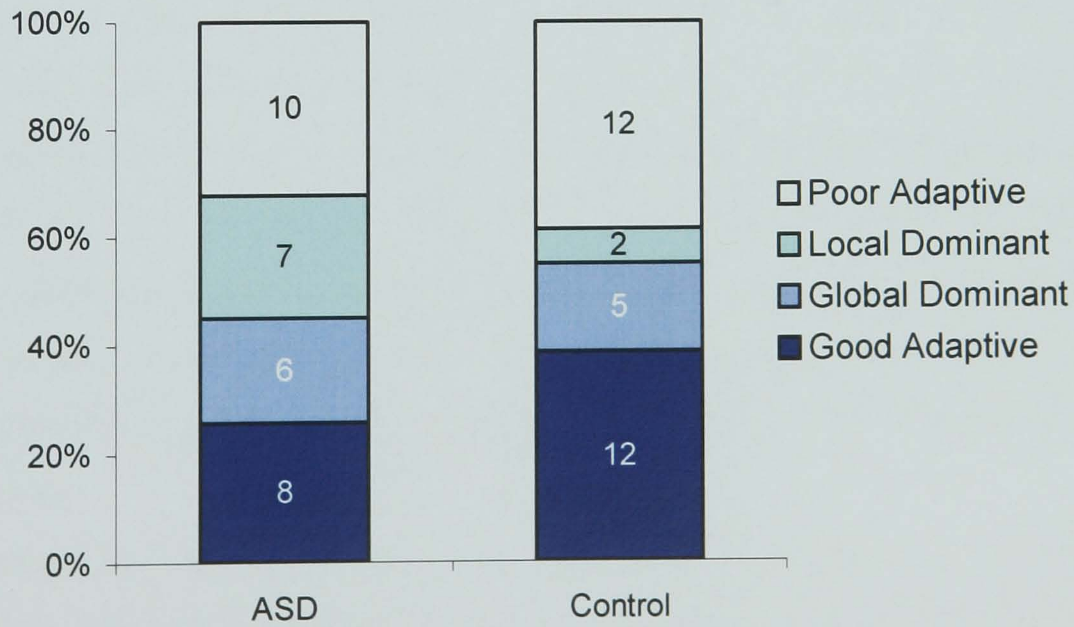


Figure 9.2 Bar chart showing the percentage (and number) of participants in each group categorised as global responders, local responders, good adaptive, and poor adaptive.

9.8 General Discussion

Chapter 9 addressed the question of whether the measures in the coherence battery were able to tap a common underlying mechanism of coherence. The results suggest that individual differences in detail-focused processing can be reliably measured in lower-level tasks across both visuo-spatial and auditory-verbal modalities, but that good local processing is not necessarily achieved at the cost of poor global processing. In the low-level visuo-spatial modality, for example, visual segmentation ability did not imply poor visual integration skills, in terms of individual differences in cognitive style.

Higher-level tasks in the visual and verbal modalities did not show simple factor solutions, suggesting they may not have tapped the same underlying cognitive bias. Higher-level tasks may be more open to strategic approaches that can largely determine performance level, and therefore the adjustment for age and IQ (made to all scores prior to PCA) may have had more impact on individual differences in these, versus lower-level, tasks.

TD participants and control individuals demonstrated the ability to adapt processing style according to task demands, suggesting a capacity to be both a good local and a good global processor. In the ASD group however, the two styles appeared to be in trade-off. The conceptualisation of central coherence processing style along a single continuum may, then, be more relevant for ASD than for individual differences in other groups.

In order to describe the pattern of individual performance, participants were characterised into subgroups: predominantly local responders, predominantly global responders, good adaptive (respond locally or globally according to which the task demanded), and poor adaptive responders. Adaptability showed clear developmental progression in the TD sample. This ability to switch local/global processing according to task demands was also related to IQ in TD and control participants, but not in ASD participants. Individuals showing a predominately global response style across tasks tended to be older in the TD sample, particularly among female participants. There was also a moderate relationship with IQ in males only, showing that those with high IQ tended to be local responders.

Interestingly, the direction of the relationship with IQ was opposite in the ASD group, such that those with low IQ tended to be local responders. It may be that the natural tendency to process for local information in ASD can only be overcome by effortful global processing. A greater proportion of participants with autism were low-functioning, while those diagnosed with Asperger syndrome were high-functioning; thus the confound between IQ and diagnosis can not be disentangled in the present sample. There was evidence of greater detail-focus in ASD versus TD and controls overall, although a predominant local bias was not pervasive in individuals with ASD. This response style was, however, more common in the ASD group than in matched controls. Relationships between coherence tasks and task composites in ASD were distinct from those in TD and control groups, and suggest that the predominant detail-focused processing style bears a different relationship to global processing in ASD. These findings are further discussed in Chapter 10, along with implications for current theories of local-global processing in ASD and TD.

Chapter 10. Discussion

10.1 Introduction

This thesis provided an in-depth investigation of detail-focused cognitive style in ASD and typical development. A central aim was to test the hypothesis that pervasive individual differences can be found in central coherence in TD children and adults, across processing modalities and levels. This was achieved by administering a battery of tasks to a large sample of TD individuals (N = 204) that spanned a wide age range (8 to 25 years). The same task battery was administered to 31 individuals with ASD and 31 age- and IQ-matched controls (9 to 21 years) in order to determine whether weak coherence is universal in individuals with ASD and correlated across levels and domains. Findings from 14 tasks from the coherence battery are presented in this thesis, from which 16 main coherence indices were selected as best indicators of local and/or global processing. The main findings are summarised below under the questions that were raised in the introductory chapters. Implications of the findings for theories of local-global processing in TD and ASD are made, followed by discussion of the limitations of the research and suggestions for the direction of future work.

10.2 Is weak central coherence universal and specific to ASD?

Demonstration of the predicted detail-focused cognitive style in ASD, compared with matched controls, reached significance in five of the 16 coherence indices (plus two further indices not selected as key measures, and three indices in low-functioning individuals only). Overall group differences occurred on three tasks where performance could either reflect local preference or poor global processing (Navon Similarity Judgment, Sentence Completion, EFT); while on at least two tasks, group differences could only reflect poor global processing (Impossible Figures, Fragmented Pictures). On these latter tasks the ASD group showed a weaker drive to integrate visual information.

At an individual level, weak central coherence was more common in participants with ASD than in matched controls, but was by no means universal. Between a quarter and two thirds of the ASD group showed weak coherence on each task, and 23% showed a consistent detail-focused cognitive style across tasks. This was in comparison to just 6% of the matched control group showing consistent detail-focus. These individual differences in coherence found within the autism spectrum most likely reflect the heterogeneity of this disorder.

Detail-focused processing appeared to be more prevalent among low-functioning than high-functioning individuals with ASD, while not showing this differentiation in the control sample. Although the number of participants in the low-functioning subgroups was small (N ≤ 6), ASD individuals with low IQ showed greater detail-focused processing than matched controls, typically when a local bias was advantageous: recognition memory for surface details (Picture Memory, Story Memory), and memory for the absolute properties of pitch (Pitch Identification). Weak central coherence was not demonstrated on these "local" coherence indices in the ASD sample as a whole. This finding may relate to diagnostic type (autism versus Asperger syndrome), however a confound between IQ and diagnosis was present in the current study, where proportionally more individuals with an autism diagnosis fell in the low IQ subgroup, than in the high IQ subgroup. A detailed-focused cognitive style may be more characteristic of autism or more characteristic of low IQ, but the samples did not include sufficient numbers to disentangle IQ and diagnosis. It remains to be seen what else distinguishes, clinically, those with consistent weak central coherence from those without.

10.3 What is the developmental course of local-global processing?

Age trends in TD were observed in most (12 of 16) coherence indices, with half of the tasks showing a progression from more detail-focused to less when a global bias was more advantageous (e.g., integrating information to form a whole on the Fragmented Pictures Task), and the remaining showing a progression from global to more detail-focused when a local bias was more advantageous (e.g., segmenting words/chords). These age trends appear to reflect changes in flexible adaptive and strategic approaches to tasks, rather than changes in cognitive style per se, but this hypothesis needs further testing in a longitudinal study. This finding is in keeping with the developmental literature, which has found that the seemingly reciprocal processes of integration and segmentation both show developmental progression. The present study found no systematic differences in the developmental trajectories for tasks that benefited from a local versus global processing style, low-level versus high-level processing, or visuo-spatial versus auditory-verbal skills. Thus reports that local processing is attained earlier in development than global processing (e.g., Burack et al., 2000; Porporino et al., 2004), were not supported by the current findings.

It was proposed that individuals might differ in their ability to use flexible and strategic approaches to tasks, perhaps independently from their natural global/local processing bias or style. In order to capture this degree of adaptability, participants were placed into four subgroups based on stable patterns of task performance: ability to adapt to the local-global task demands, and predominant local or global bias. These categorisations of individual differences were subject to developmental effects: older TD participants were characterised as "good adaptors" more often than younger TD participants. TD participants who predominantly showed a global processing style were older than those who showed a local processing style across tasks.

Age effects were not found on the coherence indices within the ASD and control groups. Furthermore, ASD and control participants categorised into the four subgroups based on local/global dominance or good/poor adaptability did not show within or between group differences in age.

10.4 What are the effects of gender on local-global processing?

Gender differences, where significant (5 of 16 coherence indices), were in general (3 of 5 indices) in the expected direction, with males more detail-focused than females. In contrast to previous literature on sex-differences, in which local versus global processing has generally been confounded with visuo-spatial versus auditory-verbal task modality, a male local bias was demonstrated on auditory-verbal tasks (Sentence Completion, Chord Sequence), as well as in visuo-spatial tasks where global bias was more advantageous (Drawing Style). The predicted male bias for local processing when detail-processing is advantageous (e.g., EFT, Un/segmented Block Design) was not demonstrated in this study, and as a consequence, the suggested (small) male superiority on visuo-spatial tasks did not contribute to the present detail-focused processing bias in males. Thus evidence was found for male-female differences in detail-focused processing, independent of general visuo-spatial ability differences. The male local bias was also more evident within high-level than low-level tasks.

In addition to gender differences in overall performance, developmental effects differed for males and females on some (7 of 16) coherence indices. Age-related improvements were more obvious in females than in males on tasks that benefited from detail-focused processing (Phoneme Segmentation, Pitch Identification, Un/segmented Block Design, EFT). In contrast, tasks in which a global processing style was more beneficial showed prominent effects of age in males, but not in females (Sentence Completion, Chord Sequence). This might fit an interpretation where developmental effects are more pronounced for attention to aspects of stimuli that come less naturally; global processing may be the default in females and local processing the preferred style in males. This finding may also correspond to gender effects found in brain maturation. For example, a cross-sectional MRI study found significant age-related increases in white matter volume of the inferior frontal gyrus in males, but not in females (Blanton et al., 2004). Regions of the corpus callosum have also been found to increase more rapidly with development in males than in females, although this did not hold after controlling for total cerebral volume (Giedd, Blumenthal, Jeffries, Rajapakse et al., 1999; Rajapakse et al., 1996). These findings are intriguing in relation to the present age and gender effects in local-global processing:

the development of white matter tracts may reflect the degree of long range connections between brain regions and relate to global processing. Thus stronger developmental effects in global processing found in males may reflect the differential development of white matter tracts between genders.

Timing of developmental trajectories also differed by gender, with the suggestion that performance improved earlier for females than males on three local coherence indices (Chord Segmentation, Un/segmented Block Design, EFT). An interesting parallel to this finding is the different developmental curves in brain development for males and females found during childhood and adolescence (Giedd, Blumenthal, Jeffries, Castellanos et al., 1999). In particular, frontal and parietal grey matter has been found to increase during pre-adolescence, peaking at around 11 years for females and 12 years for males.

10.5 What are the effects of IQ on local-global processing?

10.5.1 Typical Development

Inspection of the summary tables for each processing domain/level revealed that all eight coherence indices that benefited from a detailed processing style (e.g., chord/non-word segmentation, memory for details) showed a significant positive association with IQ in the TD sample. Furthermore, of the eight coherence indices that benefited from global processing (e.g., visual/verbal integration, global drawing style, gist recall), six were significantly and positively associated with IQ. The two exceptions were high-level verbal tasks (Homograph Reading, Sentence Completion) suggesting that the ability to process linguistic information in context for meaning did not vary systematically with IQ (including VIQ) in TD.

There was evidence that the relationship between IQ and local-global processing differed by gender. The pattern of correlations suggested qualitatively different relationships for males and females (i.e., a significant IQ correlation in one group versus no reliable correlation in the other group) on six coherence indices, although only on two indices did the difference in correlations reach statistical significance. In most cases (5 of 6 coherence indices), correlations were stronger in males between IQ and performance on tasks, whether requiring global or local processing. Thus IQ was driving performance in males more strongly than in females, when describing a scene globally (Picture Memory), cohering degraded visual information (Fragmented Pictures), retaining surface information (Story Memory), and completing block designs (less benefit from segmentation in Un/segmented Block Design). IQ was also related to reduced effect of context on homograph disambiguation (Homograph Reading Task) in males, although ceiling effects and reading ahead may explain this finding.

These gender differences may be explained by the greater variance in IQ scores in males compared to females, which was significant in the 14-16 age group (Levene's test for equality of variances: $F > 6.75$, $p < .02$ for FIQ, VIQ, FIQ-BD; see Chapter 4, Table 4.2). Alternatively, as IQ was associated with both local and global processing, this suggests that IQ may be more strongly related to a strategic executive factor in males than in females. In line with this, Duncan and colleagues (Duncan, 2005; Duncan, Emslie, Williams, Johnson, & Freer, 1996) have proposed that IQ is largely a reflection of the control functions of the frontal lobe. The IQ estimate collected from TD participants was a combination of accumulated knowledge (Information, Vocabulary) and novel problem-solving (Picture Completion, Block Design); according to Duncan the latter types of measures of fluid intelligence also tap executive functions to some extent. Although gender differences have not been examined with regard to the interrelationship between executive functions and IQ, it may be that the strategic processes in females are somehow less well tapped by the IQ tests used here.

The subgroup categorisations for stable cognitive style (local/global dominance or good/poor adaptability) were examined for IQ effects. Only the degree of adaptability, and not dominant local/global processing style, was related to IQ in the TD sample. Gender differences were indicated such that local dominant males were higher in IQ (FIQ and FIQ-BD) than global dominant males, but did not differ in age. Conversely, global dominant females were older than local dominant females, but did not differ in IQ.

10.5.2 ASD and Control groups

An association between intellectual ability and local processing demands was found in the control group, where seven of the eight "local" coherence indices correlated significantly with IQ. By contrast, only two of the eight local coherence indices showed significant correlations with IQ in the ASD group (Phoneme Segmentation, Picture Memory surface form retention). Correlations between IQ and the local coherence index were significantly stronger in the control group than in the ASD group on four measures (EFT, Pitch Identification, Story Memory surface form retention and proportion of verbatim phrase recall). Thus intellectual resources were needed to attend to local features of stimuli in control participants, but not in those with ASD.

There were no significant differences in the magnitude of the correlations between IQ and global benefiting tasks for the ASD and control groups, although the pattern of correlations differed between groups. Two global coherence indices showed a significant correlation with IQ in the ASD group, which was not found in the control group (Sentence Completion, Fragmented Pictures). Conversely, two global coherence indices did not relate to IQ in the ASD group, but did show an association in the control group, as they did in TD males (Chord Sequence, Homograph Reading).

It appears that the role of IQ in local-global processing differs in ASD from that in TD and IQ-matched controls. Local processing in particular appears to place demands on central processing resources (IQ) in TD and control participants, but less so in ASD. The IQ correlations may not have appeared in the ASD group because the low-functioning ASD participants were relatively good at tasks that benefited from a detail-focused approach. It is possible that weak coherence may accompany low-IQ in individuals with ASD. Alternatively, weak coherence may be pervasive across the majority of individuals with ASD, but is hard to detect in higher IQ samples because of compensation strategies and the ability to adapt to task demands. Local processing bias may be best captured in high-functioning individuals with ASD using tasks that are very open-ended. It was notable that in the present experiments, group differences were found mainly on tasks that required on-line processing and speeded responses. In general, however, it appears that individuals with ASD are able to obtain some understanding of the task demands, and adapt their processing approach accordingly. The natural tendency in ASD to process details is perhaps only overcome by effortful global processing (during which local focus is lost), and this ability may be dependent, in part, on IQ or some strategic executive factor.

IQ did not reliably determine subgroup (local/global dominance, good/poor adaptation) membership in the ASD group. In the control group, however, individuals with high IQ showed good adaptability. Control individuals showing good adaptability also tended to have higher IQ than those showing a predominantly global processing style.

10.6 Do stable individual differences in central coherence exist?

10.6.1 Typical Development

The main focus of this thesis was on the unity, or otherwise, of individual differences in central coherence across tasks. Although performance was driven by age, IQ, and to some extent gender, in the majority of tasks, these variables could not explain all the variance in the data. This led to the proposal that individual differences independent of age and ability could be found in the coherence measures, and suggested that cognitive style and not merely ability was being tapped.

When the effects of age, IQ, and gender were removed, stable individual differences were found in the TD sample in lower-level tasks, across and within visuo-spatial and auditory-verbal modalities. A trade-off between local and global processing was not apparent across tasks, such that ability to process details was not detrimental to proficient global processing. Correlations between higher-level tasks, as well as cross-domain and cross-level correlations, were more erratic and no clear picture emerged. This may reflect the greater difficulty of tapping cognitive style (versus strategy/ability) in higher-level tasks or a genuine lack of unity among local-global biases at these different levels and across modality. It is therefore possible that processing bias was tapped more clearly by lower-level tasks. The wider implication of these findings for models of local and global processing includes the suggestion that distinct mechanisms may need to be found underlying individual differences in featural processing and in configural processing, with the two mechanisms not operating in trade-off. To the extent that a common mechanism underlies manifestations of weak coherence, this may operate at a relatively low perceptual level.

As good global processing did not compromise local processing, this suggests that many TD individuals were able to adapt processing style according to task demands. This finding may relate to the field dependence/independence literature where the ability to adapt processing style may reflect the ability to overcome and switch attention from the gestalt, rather than a natural tendency not to see the gestalt (weak coherence). This degree of adaptability or overcoming of a natural bias may also relate to aspects of higher-order cognition, such as executive functions. While weak coherence itself does not seem to be a mere side effect of poor executive functions in ASD (Booth et al., 2003; Pellicano, Maybery et al., 2005; Teunisse et al., 2001), the ability to flexibly resist this natural bias may depend on executive function/IQ capacities at least in TD.

It was proposed that individuals could be characterised according to their performance across the whole coherence battery, as showing a predominant local/global processing style, or good/poor adaptability to task demands. Longitudinal studies may determine whether these individual differences (subgroups) are stable with age. It is speculated that TD individuals may shift in development from a predominant local or global processing style to show good adaptive processing, with this occurring perhaps after a certain age or at a certain IQ or MA level.

10.6.2 ASD and Control groups

Overall, the results suggest that in ASD the predominant detail-focused processing style bears a different relationship to global processing than in TD or IQ-matched controls. While TD and control individuals showed patterns of task relations suggestive of an ability to be both a good local and a good global processor, in ASD the two processing styles seemed to be in trade-off as shown by the pattern of correlations.

As in the TD sample, ASD and control participants were grouped by whether they demonstrated a predominately local or global response style, or whether they showed good or poor adaptability to task demands. A dominant local processing style was not present in all individuals with ASD, although a greater proportion (23%) showed this cognitive style compared to the control group (6%). The number of participants classified as showing a global dominant processing style, or good/poor adaptability, did not differ between ASD and control groups. Thus individuals with ASD did not show a specific deficit in the ability to adapt processing style to task demands. Again, there is a need for further studies to see if subgroup membership is stable with development and what clinical factors are associated.

Together, the findings do suggest a top-down influence on the central coherence model. First, adaptability to task demands was strongly associated with IQ in the TD sample, but not in the ASD group. Secondly, there was greater evidence of a trade-off between segmentation and integration in the ASD group than compared to the TD sample. As Happé and Frith (2006) propose, several features of autism can be characterised by reduced top-down modulation of executive control, with local processing being the default when this control is weak or absent. Thus, there may be important theoretical commonalities between the executive function and central coherence accounts.

10.7 How do the results compare with previous studies?

The hypothesis that central coherence is a unitary concept, pervasive across different tasks, has been tested in a growing number of studies, but without conclusive answers to date. Previous studies have, in general, tested the pervasiveness of central coherence within the same processing modality. This runs the risk of assessing a specific ability (e.g., general visuo-spatial ability), or another dominant processing style (e.g., the verbaliser-visualiser dichotomy; A. Richardson, 1977), rather than central coherence. Greater convergent validity of central coherence would be demonstrated by cross-domain and cross-level associations. Although the findings are mixed, evidence of cross-modality coherence has been found in the ASD literature (Loth et al., 2006; Teunisse et al., 2001), with the suggestion that a trade-off in local-global processing is more apparent in ASD than in individuals with TD or MLD. The present findings support this account, in particular within the low-level visuo-spatial and auditory processing domains.

Pellicano and colleagues (Pellicano, 2004; Pellicano, Maybery et al., 2005) found performance across visuo-spatial tasks that benefited from a local processing style was associated within ASD and TD, but was also positively associated with performance on tasks requiring global integrative ability. Although this goes against a unitary concept of central coherence, this finding can be interpreted as the capacity for an individual to be both a good local processor and a good global processor, within the framework proposed in this thesis. After removing the effects of age and ability, Pellicano found a positive association between local and global processing to remain between two tasks in TD and control participants, but not significantly in ASD. Similar to the findings in the present study within the low-level visuo-spatial domain, this hints that a trade-off between local and global processing may be more apparent in ASD, compared to controls.

It is clear that the autism spectrum is heterogeneous and this may account for some discrepancies in research findings. The present study found approximately a quarter of individuals with ASD to have a predominantly local bias, with a third showing poor adaptability. This might explain why predictions of weak central coherence have not necessarily been fulfilled within previous studies of ASD versus controls.

10.8 Can the results help us choose between alternative theoretical accounts of weak central coherence?

Common to the two main alternative theoretical accounts of weak coherence findings put forward by Plaisted and by Mottron (see Chapter 2, Section 2.5) is that, while perception and discrimination of local properties is enhanced, global processing skills are intact. Both views also situate local superiority at very low levels, with these perceptual abnormalities having downstream effects on high-level processes. The subgroup analysis found an equivalent proportion of ASD and control participants were categorised as showing a predominant global bias. This finding may provide evidence for intact global processing in ASD, however general support for these alternative accounts was not found in the present study. ASD and control group differences, where apparent, were at least as evident in lack of global processing (e.g., poor visual integration on Impossible Figures and Fragmented Pictures) as in good local processing (e.g., proficiency on EFT, local bias on Navon Similarity Judgment and Sentence Completion). Furthermore, low-level perceptual tasks were not more likely to capture local processing in ASD than high-level tasks, despite the finding that low-level tasks showed greater unity than high-level tasks. The conceptualisation of weak central coherence as a bias towards featural processing, even on tasks where taking a global approach would have been advantageous, was broadly supported by the present findings. Superior local processing was found in individuals with ASD and low IQ, however, and warrants further investigation in larger samples.

The tasks in the present study were not designed to address Baron-Cohen's empathising-systemising account of ASD. The suggested facility for systemising in ASD, proposed to involve attention to exact detail, may be supported by the finding of superior EFT performance in ASD. However, the empathising-systemising account does not predict deficits in global processing (but rather an ability to integrate components of a system) and thus findings of reduced global processing in the present study do not support this theory. The slight male local bias found in the TD sample lends some support to Baron-Cohen's extreme male brain theory of autism. However, indictors of "maleness" (i.e., testosterone level, digit ratio), which may vary within each biological gender, were not measured in the present study and may have provided more definitive information on whether a local bias may be linked to a male brain.

10.9 Limitations of the study

This study was pioneering in its attempt to provide a comprehensive assessment of central coherence across and within different processing modalities and levels. This allowed for an individual differences approach to the examination of the pervasiveness of detail-focused processing style in individuals with and without ASD. The research did, however, suffer from some limitations in design and sample selection, and these are discussed below.

Despite the testing of a large sample of TD individuals, with equivalent numbers of males and females, the IQ distribution was not even across age group and gender. Overall, males had higher FIQ (but not FIQ-BD) than females. The IQ profile across age group was also flatter in females than in males (where the youngest and oldest age groups had higher FIQ scores). Although this variability could be taken into account on some tasks by covariation, it may still limit some conclusions made regarding gender effects.

Individuals with autism and Asperger syndrome were not evenly represented within high- and low-functioning subgroups, thus a confound existed between diagnosis and IQ. In addition, the group with low IQ was small. The main reason for this was the greater practicality of testing high-functioning individuals in this study. As the focus was on whether central coherence was a unitary construct, this necessitated the administration of a large task battery. The attentional demands required limited the accessibility of the battery for low-functioning participants (who, within the autism spectrum, typically have a diagnosis of autism). The task battery was very demanding for the low-functioning participants in this study and required many more testing sessions than in the high-functioning group.

Low-functioning individuals were not initially considered for inclusion in the study, particularly as high-functioning individuals with ASD are traditionally considered to represent a "pure" form of autism (Rutter, 1983). However the low-functioning individuals with ASD who were amenable to testing were retained in the data set as, somewhat surprisingly, they were well able to complete many of the coherence tasks. As it turned out, the results suggested detail focus might be more clearly seen in these individuals making inclusion of more low-functioning participants in future studies a worthwhile aim. As discussed above, high-functioning individuals may 'second-guess' the underlying aim of tasks or use compensatory strategies that may mask cognitive style.

A limitation of the study was that not all coherence measures were able to discriminate ASD from controls. Consideration needs to be made of the fact that several tasks were newly designed and may not have been sensitive to detecting weak central coherence despite extensive piloting. Although contrary to weak coherence theory, some findings were consistent with the ASD literature (e.g., intact global processing of musical stimuli). A slightly different design may have resolved some conflicting findings (e.g., presenting the target note before versus after chord presentation on the Chord Segmentation Task), although on some tasks it was unclear why the predicted results were not found (e.g., verbatim phrase recall on Story Memory, segmentation ability on Phoneme Segmentation).

Another limitation to task design was the inability to equate coherence tasks on their level of difficulty or discriminability, across low and high processing levels, and across visual and verbal modalities. Although variation by age and ability is provided by the TD sample, it would be difficult to gauge whether, overall, the auditory-processing tasks were more difficult than the visuo-spatial tasks, for example. Lack of equivalence in task difficulty across processing modalities/levels, and the difficulty in providing a metric to do so, may provide an explanation for the inconsistencies in central coherence found within- and between-domains. It must be noted that this is a pervasive challenge in cognitive research.

10.10 Directions for future research

One aim for future work is to examine the cognitive phenotype across the full autism spectrum. It would be interesting to target specifically the lower ability range, given that local bias appeared to be more characteristic of this group. ASD is well acknowledged as a heterogeneous disorder at every level and Szatmari et al. (2002) have suggested that low-functioning autism might be genetically distinct from high-functioning autism. Processing style might potentially provide important information on the distinctiveness of autism and Asperger syndrome, or low- and high-functioning ASD, and could be examined in relation to genetic subgrouping.

The nature of the subgroups defined by local/global dominance and good/poor adaptability warrants further investigation. In particular, it would be interesting to establish whether the subgroups can be distinguished by tests of executive functioning. This may inform whether degree of adaptability can be considered under the umbrella term of executive functions (and specific sub-functions such as set-shifting or generativity), or whether these processes are distinct. Future studies need to address the developmental progression of local-global processing alongside the ability to adapt to task demands.

Current literature outside the field of autism suggests that negative mood may be correlated with, and perhaps causally related to, local-processing bias in healthy and depressed participants (Derryberry & Tucker, 1994; Gasper & Clore, 2002; Hesse & Spies, 1996). This could question the main hypothesis of this thesis, that central coherence can be regarded as a stable cognitive style. Since depression and anxiety are common in ASD (e.g., Gadow, DeVincent, Pomeroy, & Azizian, 2004; Ghaziuddin, 2002) it is important to test the possible hypothesis that negative mood, rather than ASD itself, is responsible for weak coherence findings. Future work with the current data set will examine whether state/trait anxiety and depression measures taken at the time of testing show any association to an individual's detail-focused processing style.

To provide some ecological validity to a study, it is important to test the association between cognitive abilities/deficits and behavioural symptoms. A self-report questionnaire was given to all study participants, asking them about their preference for detail in everyday life (e.g., "Do you notice little things that other people miss?"), and about the converse tendency to see patterns and process information in context for meaning. Future work will examine the relationship between real-life examples of weak central coherence, as assessed by this questionnaire, and experimental measures.

10.11 Concluding comment

In conclusion, the work presented in this thesis may extend the original theory of weak central coherence as proposed by Frith (1989), and Frith and Happé (1994). The main conclusion made from the TD and control data is that an individual's tendency for local processing and for global processing is not necessarily in trade-off. There is a suggestion that these processes are more in trade-off in ASD and that a local processing style is more typical and less costly in terms of processing resources. In order to characterise stable individual differences in central coherence it appeared necessary to take into account an adaptability or strategic executive factor, which was not accounted for by age or IQ. The conceptualisation of weak central coherence as a natural bias within ASD, which may be overcome by effortful global processing, appears to hold in the present data although this style is not universal in individuals with ASD. More work is required to clarify the way in which this description at the cognitive level may map onto the neural and behavioural levels.

Weak central coherence theory, along with the influential deficit accounts of ASD (theory of mind, executive function), has contributed to a better insight into information-processing by people with ASD. It is hoped that this thesis can contribute to solving the enigma of how individuals with ASD think and learn. Differences in cognitive style and weak coherence are now beginning to be recognised within educational approaches (e.g., Jacobsen, 2005). The implication that ASD may not be defined solely in terms of deficits, but also by strengths in detail-focused processing can only be encouraging to individuals with ASD, their families, and educators.

References

Aginsky, V., & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition, 7*(1-3), 147-162.

Akshoomoff, N. A., Feroleto, C. C., Doyle, R. E., & Stiles, J. (2002). The impact of early unilateral brain injury on perceptual organization and visual memory. *Neuropsychologia, 40*(5), 539-561.

Akshoomoff, N. A., & Stiles, J. (1995a). Developmental trends in visuospatial analysis and planning: I. Copying a complex figure. *Neuropsychology, 9*(3), 364-377.

Akshoomoff, N. A., & Stiles, J. (1995b). Developmental trends in visuospatial analysis and planning: II. Memory for a complex figure. *Neuropsychology, 9*(3), 378-389.

Akshoomoff, N. A., & Stiles, J. (1996). The influence of pattern type on children's block design performance. *Journal of the International Neuropsychological Society, 2*(5), 392-402.

Ameli, R., Courchesne, E., Lincoln, A., Kaufman, A. S., & Grillon, C. (1988). Visual memory processes in high-functioning individuals with autism. *Journal of Autism and Developmental Disorders, 18*(4), 601-615.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, M. (2001). Conceptions of intelligence. *Journal of Child Psychology and Psychiatry, 42*(3), 287-298.

Applebaum, E., Egel, A. L., Koegel, R. L., & Imhoff, B. (1979). Measuring musical abilities of autistic children. *Journal of Autism and Developmental Disorders, 9*(3), 279-285.

Arceneaux, J. M., Cheramie, G. M., & Smith, C. W. (1996). Gender differences in WAIS-R age-corrected scaled scores. *Perceptual and Motor Skills, 83*(3), 1211-1215.

Aslin, R. N., & Smith, L. B. (1988). Perceptual development. *Annual Review of Psychology, 39*, 435-473.

Aurnhammer-Frith, U. (1969). Emphasis and meaning in recall in normal and autistic children. *Language and Speech, 12*(1), 29-38.

Baharloo, S., Service, S. K., Risch, N., Gitschier, J., & Freimer, N. B. (2000). Familial aggregation of absolute pitch. *American Journal of Human Genetics, 67*(3), 755-758.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., et al. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine, 25*(1), 63-77.

Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders, 28*(5), 369-392.

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., et al. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(6), 694-702.

Baker, P., Piven, J., & Sato, Y. (1998). Autism and tuberous sclerosis complex: Prevalence and clinical features. *Journal of Autism and Developmental Disorders, 28*(4), 279-285.

Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., & Reiss, A. L. (2004). White matter structure in autism: Preliminary evidence from diffusion tensor imaging. *Biological Psychiatry, 55*(3), 323-326.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254.

Baron-Cohen, S., Bolton, P., Wheelwright, S., Scahill, V., Short, L., Mead, G., et al. (1998). Autism occurs more often in families of physicists, engineers, and mathematicians. *Autism, 2*, 296-301.

Baron-Cohen, S., & Hammer, J. (1997). Parents of children with Asperger syndrome: What is the cognitive phenotype? *Journal of Cognitive Neuroscience, 9*(4), 548-554.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry, 38*(7), 813-822.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1430), 361-374.

Baron-Cohen, S., Ring, H., Wheelwright, S., Bullmore, E., Brammer, M., Simmons, A., et al. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience, 11*(6), 1891-1898.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience.* New York: Oxford University Press.

Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders, 5*(1), 47-78.

Baron-Cohen, S., Wheelwright, S., Stott, C., Bolton, P., & Goodyer, I. (1997). Is there a link between engineering and autism? *Autism, 1*(1), 101-109.

Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific developmental receptive language disorder: I. The children. *British Journal of Psychiatry, 126*, 127-145.

Bartlett, F. (1932). *Remembering.* Cambridge, England: Cambridge University Press.

Bartolucci, G., Pierce, S., Streiner, D., & Eppel, P. T. (1976). Phonological investigation of verbal autistic and mentally retarded subjects. *Journal of Autism and Childhood Schizophrenia, 6*(4), 303-316.

Bartolucci, G., & Pierce, S. J. (1977). A preliminary comparison of phonological development in autistic, normal, and mentally retarded subjects. *British Journal of Disorders of Communication, 12*(2), 137-147.

Basso, M. R., Schefft, B. K., Ris, M., & Dember, W. N. (1996). Mood and global-local visual processing. *Journal of the International Neuropsychological Society, 2*(3), 249-255.

Bauman, M. L., & Kemper, T. L. (2005). Neuroanatomic observations of the brain in autism: A review and future directions. *International Journal of Developmental Neuroscience, 23*(2-3), 183-187.

Beech, H., & Graham, M. (1967). Note on use of sentence completion in assessing overt aggression in normal school children. *Psychological Reports, 20*(1), 9-10.

Beery, K. E. (1997). *Developmental test of visual-motor integration* (4th ed.). New Jersey: Modern Curriculum Press.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St. George, M. (2000). The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12*(Suppl 1), 7-29.

Belmonte, M. K., Allen, G., Beckel-Mitchener, A., Boulanger, L. M., Carper, R. A., & Webb, S. J. (2004). Autism and abnormal development of brain connectivity. *Journal of Neuroscience, 24*(42), 9228-9231.

Benowitz, L. I., Moya, K. L., & Levine, D. N. (1990). Impaired verbal reasoning and constructional apraxia in subjects with right hemisphere damage. *Neuropsychologia, 28*(3), 231-241.

Bentley, A. (1985). *Measures of musical abilities.* Windsor: NFER-NELSON.

Berger, H. J., Aerts, F. H., van Spaendonck, K. P., Cools, A. R., & Teunisse, J.-P. (2003). Central coherence and cognitive shifting in relation to social improvement in high-functioning young adults with Autism. *Journal of Clinical and Experimental Neuropsychology, 25*(4), 502-511.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2003). Motion perception in autism: A "complex" issue. *Journal of Cognitive Neuroscience, 15*(2), 218-225.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry, 175*(5), 444-451.

Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and nonmusicians. *Science, 185*(4150), 537-539.

Beversdorf, D. Q., Smith, B. W., Crucian, G. P., Anderson, J. M., Keillor, J. M., Barrett, A. M., et al. (2000). Increased discrimination of "false memories" in autism spectrum disorder. *Proceedings of the National Academy of Sciences of the United States of America, 97*(15), 8734-8737.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5*(1), 1-30.

Bhatt, R. S., Rovee-Collier, C., & Shyi, G. C. (1994). Global and local processing of incidental information and memory retrieval at 6 months. *Journal of Experimental Child Psychology, 57*(2), 141-162.

Biederman, I. (1972). Perceiving real-world scenes. *Science, 177*(4043), 77-80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*(2), 143-177.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception and Psychophysics, 59*(7), 1098-1107.

Bihrle, A. M., Bellugi, U., Delis, D. C., & Marks, S. (1989). Seeing either the forest or the trees: Dissociation in visuospatial processing. *Brain and Cognition, 11*(1), 37-49.

Bird, J., & Bennett, A. F. (1974). A developmental study of recognition of pictures and nouns. *Journal of Experimental Child Psychology, 18*(1), 117-126.

Bishop, D. V. M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children.* Hove, England: Psychology Press.

Bjorklund, D. F., & Jacobs, J. W. (1985). Associative and categorical processes in children's memory: The role of automaticity in the development of organization in free recall. *Journal of Experimental Child Psychology, 39*(3), 599-617.

Blanton, R. E., Levitt, J. G., Peterson, J. R., Fadale, D., Sporty, M. L., Lee, M., et al. (2004). Gender differences in the left inferior frontal gyrus in normal children. *Neuroimage, 22*, 626-636.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory and Cognition, 8*(6), 631-642.

Bolton, P., Murphy, M., Macdonald, H., & Whitlock, B. (1997). Obstetric complications in autism: Consequences or causes of the condition? *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(2), 272-281.

Bolton, P., & Rutter, M. (1990). Genetic influences in autism. *International Review of Psychiatry, 2*(1), 67-80.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience, 15*(2), 226-235.

Booth, R., Charlton, R., Hughes, C., & Happé, F. (2003). Disentangling weak coherence and executive dysfunction: planning drawing in autism and attention-deficit/hyperactivity disorder. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 358*(1430), 387-392.

Borges, M. A., Stepnowsky, M. A., & Holt, L. H. (1977). Recall and recognition of words and pictures by adults and children. *Bulletin of the Psychonomic Society, 9*(2), 113-114.

Bowler, D. M., Gardiner, J. M., Grice, S., & Saavalainen, P. (2000). Memory illusions: False recall and recognition in adults with Asperger's syndrome. *Journal of Abnormal Psychology, 109*(4), 663-672.

Bowler, D. M., Matthews, N. J., & Gardiner, J. M. (1997). Asperger's syndrome and memory: Similarity to autism but not amnesia. *Neuropsychologia, 35*(1), 65-70.

Bradford, Y., Haines, J., Hutcheson, H., Gardiner, M., Braun, T., Sheffield, V., Cassavant, T., Huang, W., Wang, K., Vieland, V., Folstein, S., Santangelo, S., Piven, J. (2001). Incorporating language phenotypes strengthens evidence of linkage to autism. *American Journal of Medical Genetics, 105*(6), 539-547.

Brainerd, C. J., & Gordon, L. L. (1994). Development of verbatim and gist memory for numbers. *Developmental Psychology, 30*(2), 163-177.

Brainerd, C. J., & Mojardin, A. (1998). Children's and adults' spontaneous false memories: Long-term persistence and mere-testing effects. *Child Development, 69*(5), 1361-1377.

Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24*(4), 396-439.

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development, 73*(5), 1363-1377.

Brambilla, P., Hardan, A., di Nemi, S. U., Perez, J., Soares, J. C., & Barale, F. (2003). Brain anatomy and development in autism: Review of structural MRI studies. *Brain Research Bulletin, 61*(6), 557-569.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*(4), 331-350.

Brian, J. A., & Bryson, S. E. (1996). Disembedding performance and recognition memory in autism/PDD. *Journal of Child Psychology and Psychiatry, 37*(7), 865-872.

Briskman, J., Happé, F., & Frith, U. (2001). Exploring the cognitive phenotype of autism: Weak "central coherence" in parents and siblings of children in autism: II. Real-life skills and preferences. *Journal of Child Psychology and Psychiatry, 42*(3), 309-316.

Brock, J., Brown, C. C., Boucher, J., & Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Development and Psychopathology, 14*(2), 209-224.

Brosnan, M., Demetre, J., Hamill, S., Robson, K., Shepherd, H., & Cody, G. (2002). Executive functioning adults and children with developmental dyslexia. *Neuropsychologia, 40*(12), 2144-2155.

Brosnan, M. J., Scott, F. J., Fox, S., & Pye, J. (2004). Gestalt processing in autism: Failure to process perceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry, 45*(3), 459-469.

Brown, A. L. (1973). Judgments of recency for long sequences of pictures: The absence of a developmental trend. *Journal of Experimental Child Psychology, 15*(3), 473-480.

Bruyer, R., & Scailquin, J.-C. (2000). The fate of global precedence with age. *Experimental Aging Research, 26*(4), 285-314.

Bruyer, R., Scailquin, J.-C., & Samson, D. (2003). Aging and the locus of the global precedence effect: A short review and new empirical data. *Experimental Aging Research, 29*(3), 237-368.

Burack, J. A. (1994). Selective attention deficits in persons with autism: preliminary evidence of an inefficient attentional lens. *Journal of Abnormal Psychology, 103*(3), 535-543.

Burack, J. A., Enns, J. T., Iarocci, G., & Randolph, B. (2000). Age differences in visual search for compound patterns: Long- versus short-range grouping. *Developmental Psychology, 36*(6), 731-740.

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34*(4), 263-272.

Burnette, C. P., Mundy, P. C., Meyer, J. A., Sutton, S. K., Vaughan, A. E., & Charak, D. (2005). Weak central coherence and its relations to theory of mind and anxiety in autism. *Journal of Autism and Developmental Disorders, 35*(1), 63-73.

Cairns, E., Malone, S., Johnston, J., & Cammock, T. (1985). Sex differences in Children's Group Embedded Figures Test performance. *Personality and Individual Differences, 6*(5), 653-654.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*(4275), 312-314.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge UK: Cambridge University Press.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(8), 1839-1849.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Journal of the American Medical Association, 285*(24), 3093-3099.

Charman, T., & Baron-Cohen, S. (1993). Drawing development in autism: The intellectual to visual realism shift. *British Journal of Developmental Psychology, 11*(2), 171-185.

Chen, S., & Levi, D. M. (1996). Angle judgment: Is the whole the sum of its parts? *Vision Research, 36*(12), 1721-1735.

Chen, Y., Nakayama, K., Levy, D., Matthysse, S., & Holzman, P. (2003). Processing of global, but not local, motion direction is deficient in schizophrenia. *Schizophrenia Research, 61*(2-3), 215-227.

Chess, S., Fernandez, P., & Korn, S. (1978). Behavioral consequences of congenital rubella. *Journal of Pediatrics, 93*(4), 699-703.

Chipman, K., & Kimura, D. (1998). An investigation of sex differences on incidental memory for verbal and pictorial material. *Learning and Individual Differences, 10*(4), 259-272.

Chung, M. K., Dalton, K. M., Alexander, A. L., & Davidson, R. J. (2004). Less white matter concentration in autism: 2D voxel-based morphometry. *Neuroimage, 23*(1), 242-251.

Coates, S. (1972). *Preschool Embedded Figures Tests: PEFT.* Palo Alto, CA: Consulting Psychologists Press.

Cobrinik, L. (1977). Use of a sentence completion method for the study of language of severely disturbed children. *Acta Paedopsychiatrica, 42*(5), 170-178.

Cody, H., Pelphrey, K., & Piven, J. (2002). Structural and functional magnetic resonance imaging of autism. *International Journal of Developmental Neuroscience, 20*(3-5), 421-438.

Cohen, I. L. (1994). An artificial neural-network analog of learning in autism. *Biological Psychiatry, 36*(1), 5-20.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, L. B., & Younger, B. A. (1984). Infant perception of angular relations. *Infant Behavior and Development, 7*(1), 37-47.

Cohen, M. J. (1997). *Children's Memory Scale.* London: The Psychological Corporation.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*, 337-367.

Colombo, J., Mitchell, D., Coldren, J. T., & Freeseman, L. J. (1991). Individual differences in infant visual attention: Are short lookers faster processors or feature processors? *Child Development, 62*(6), 1247-1257.

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. In J. Colombo & J. W. Fagen (Eds.), *Individual differences in infancy: Reliability, stability, prediction* (pp. 193-227). Hillsdale, NJ: Lawrence Erlbaum.

Comalli, P. E. (1965). Cognitive functioning in a group of 80-90-year-old men. *Journal of Gerontology, 29*, 9-13.

Corah, N. L., & Gospodinoff, E. J. (1966). Color-form and whole-part perception in children. *Child Development, 37*(4), 837-842.

Costa-Giomi, E. (2003). Young children's harmonic perception. *Annals of the New York Academy of Sciences, 999*(1), 477-484.

Costa-Giomi, E., Gilmour, R., Siddell, J., & Lefebvre, E. (2001). Absolute pitch, early musical instruction, and spatial abilities. *Annals of the New York Academy of Sciences, 930*(1), 394-396.

Courchesne, E., Karns, C., Davis, H., Ziccardi, R., Carper, R., Tigue, Z., et al. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology, 57*(2), 245-254.

Cowie, R., & Perrott, R. (1993). From line drawings to impressions of 3D objects: developing a model to account for the shapes that people see. *Image and Vision Computing, 11*(6), 342-352.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology, 19*(2), 141-177.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65*(2), 171-237.

Cyr, J., & Brooker, B. H. (1984). Use of appropriate formulas for selecting WAIS--R short forms. *Journal of Consulting and Clinical Psychology, 52*(5), 903-905.

Dawson, G., Webb, S., Schellenberg, G. D., Dager, S., Friedman, S., Aylward, E., et al. (2002). Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology, 14*(3), 581-611.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., et al. (2003). Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology, 25*(2), 242-254.

Delis, D. C., Kiefner, M. G., & Fridlund, A. J. (1988). Visuospatial dysfunction following unilateral brain damage: Dissociations in hierarchical hemispatial analysis. *Journal of Clinical and Experimental Neuropsychology, 10*(4), 421-431.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Leaning Test manual.* San Antonio, TX: The Psychological Corporation.

Demsky, Y., Carone, D. A., Jr., Burns, W. J., & Sellers, A. (2000). Assessment of visual-motor coordination in 6- to 11-yr.-olds. *Perceptual and Motor Skills, 91*(1), 311-321.

Deregowski, J. (1969). Perception of the two-pronged trident by two- and three-dimensional perceivers. *Journal of Experimental Psychology, 82*(1), 9-13.

Derryberry, D., & Reed, M. A. (1998). Anxiety and attentional focusing: Trait, state and hemispheric influences. *Personality and Individual Differences, 25*(4), 745-761.

Derryberry, D., & Tucker, D. M. (1994). Motivating the focus of attention. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 167-196). San Diego, CA: Academic Press.

Dixon, P., & Twilley, L. C. (1999). An integrated model of meaning and sense activation and disambiguation. *Brain and Language, 68*(1-2), 165-171.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121*(3), 275-277.

Donnelly, N., Found, A., & Muller, H. J. (1999). Searching for impossible objects: Processing form and attributes in early vision. *Perception and Psychophysics, 61*(4), 675-690.

Dukette, D., & Stiles, J. (1996). Children's analysis of hierarchical patterns: Evidence from a similarity judgment task. *Journal of Experimental Child Psychology, 63*(1), 103-140.

Dukette, D., & Stiles, J. (2001). The effects of stimulus density on children's analysis of hierarchical patterns. *Developmental Science, 4*(2), 233-251.

Dulaney, C. L., Marks, W., & Devine, C. (1994). Global/local processing in mentally retarded and nonretarded persons. *Intelligence, 19*(2), 245-261.

Duncan, J. (2005). Frontal lobe function and general intelligence: Why it matters. *Cortex, 41*(2), 215-217.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*(3), 257-303.

Elkind, D., Anagnostopoulou, R., & Malone, S. (1970). Determinants of part-whole perception in children. *Child Development, 41*(2), 391-397.

Elkind, D., Koegler, R., & Go, E. (1964). Studies in perceptual development: II. Part-whole perception. *Child Development, 35*(1), 81-90.

Elliot, C. D. (1990). *Differential Abilities Scale.* San Antonio, TX: Psychological Corporation.

Elliot, C. D., Murray, D. J., & Pearson, L. S. (1996). *British Ability Scales.* Windsor: NFER Nelson.

Enns, J. T., & Girgus, J. S. (1985). Developmental changes in selective and integrative visual attention. *Journal of Experimental Child Psychology, 40*(2), 319-337.

Enns, J. T., & Kingstone, A. (1995). Access to global and local properties in visual search for compound stimuli. *Psychological Science, 6*(5), 283-291.

Fagot, J., & Deruelle, C. (1997). Processing of global and local visual information and hemispheric specialization in humans (Homo sapiens) and baboons (Papio papio). *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 429-442.

Fantz, R. L. (1961). The origin of form perception. *Scientific American, 204*(5), 66-72.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science, 140*(3564), 296-297.

Fein, D., Lucci, D., & Waterhouse, L. (1990). Brief report: Fragmented drawings in autistic children. *Journal of Autism and Developmental Disorders, 20*(2), 263-269.

Ferman, T. J., Primeau, M., Delis, D. C., & Jampala, C. V. (1999). Global-local processing in schizophrenia: Hemispheric asymmetry and symptom-specific interference. *Journal of the International Neuropsychological Society, 5*(5), 442-451.

Ferstl, E. C., & von Cramon, D. (2001). The role of coherence and cohesion in text comprehension: An event-related fMRI study. *Cognitive Brain Research, 11*(3), 325-340.

Fink, G., Halligan, P. W., Marshall, J. C., Frith, C., Frackowiak, R. S., & Dolan, R. J. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain, 120*(10), 1779-1791.

Fink, G., Marshall, J., Halligan, P., & Dolan, R. (1999). Hemispheric asymmetries in global/local processing are modulated by perceptual salience. *Neuropsychologia, 37*(1), 31-40.

Fisher, N. C. (2002). *The relationship of theory of mind to language and executive function in children with autism and children with moderate learning difficulties.* Unpublished Ph.D. Thesis, King's College, University of London, London.

Fodor, J., & Bever, T. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*(5), 414-420.

Fogliani-Messina, T. M., Fogliani, A. M., & di Nuovo, S. (1983). Embedded Figures Test in old age: A psychometric note. *Perceptual and Motor Skills, 56*(1), 284-286.

Fombonne, E. (2003). The prevalence of autism. *Journal of the American Medical Association, 289*(1), 87-89.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18*(4), 281-294.

Fox, R., & Oross, S. (1990). Mental retardation and perception of global motion. *Perception and Psychophysics, 48*(3), 252-258.

Foxton, J. M., Stewart, M. E., Barnard, L., Rodgers, J., Young, A. H., O'Brien, G., et al. (2003). Absence of auditory 'global interference' in autism. *Brain, 126*(12), 2703-2709.

Freeseman, L. J., Colombo, J., & Coldren, J. T. (1993). Individual differences in infant visual attention: Four-month-olds' discrimination and generalization of global and local stimulus properties. *Child Development, 64*(4), 1191-1203.

Frick, J. E., Colombo, J., & Allen, J. R. (2000). Temporal sequence of global-local processing in 3-month-old infants. *Infancy, 1*(3), 375-386.

Friston, K. J. (1999). Schizophrenia and the disconnection hypothesis. *Acta Psychiatrica Scandinavica Supplementum, 99*(Suppl 395), 68-79.

Frith, C. (2004). Is autism a disconnection disorder? *The Lancet Neurology, 3*(10), 577.

Frith, U. (1989). *Autism: Explaining the enigma.* Oxford, UK: Blackwell Publishing.

Frith, U. (1991). *Autism and Asperger syndrome.* New York: Cambridge University Press.

Frith, U. (2001). Mind blindness and the brain in Autism. *Neuron, 32*(6), 969-979.

Frith, U. (2003). *Autism: Explaining the enigma* (2nd ed.). Oxford, UK: Blackwell Publishing.

Frith, U. (2004). Emanuel Miller lecture: Confusions and controversies about Asperger syndrome. *Journal of Child Psychology and Psychiatry, 45*(4), 672-686.

Frith, U., & Baron-Cohen, S. (1987). Perception in autistic children. In D. Cohen, A. Donnellan & R. Paul (Eds.), *Handbook of Autism and Disorders of Atypical Development.* (pp. 85-102). New York: Wiley.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition, 50*(1-3), 115-132.

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology, 1*(4), 329-342.

Gadow, K. D., DeVincent, C. J., Pomeroy, J., & Azizian, A. (2004). Psychiatric symptoms in preschool children with PDD and clinic and comparison samples. *Journal of Autism and Developmental Disorders, 34*(4), 379-393.

Gallagher, A. (1995). The development of a phonological assessment battery: Research background. *Educational and Child Psychology, 12*(1), 18-24.

Gasper, K. (2004). Do you see what I see? Affect and visual information processing. *Cognition and Emotion, 18*(3), 405-421.

Gasper, K., & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information. *Psychological Science, 13*(1), 34-40.

Gepner, B., Mestre, D., Masson, G., & de Schonen, S. (1995). Postural effects of motion vision in young autistic children. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 6*(8), 1211-1214.

Gepner, B., & Mestre, D. R. (2002). Brief report: Postural reactivity to fast visual motion differentiates autistic from children with Asperger Syndrome. *Journal of Autism and Developmental Disorders, 32*(3), 231-238.

Gerland, G. (1997). *A real person: Life on the outside* (J. Tate, Trans.). London: Souvenir Press.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17*(3), 324-363.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science, 4*(5), 294-298.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(2), 245-262.

Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Review of Psychology, 54*, 91-114.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430-445.

Geurts, H. M., Verte, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry, 45*(4), 836-854.

Ghaziuddin, M. (2002). Asperger syndrome: Associated psychiatric and medical conditions. *Focus on Autism and Other Developmental Disabilities, 17*(3), 138-144.

Ghim, H.-R., & Eimas, P. D. (1988). Global and local processing by 3- and 4-month-old infants. *Perception and Psychophysics, 43*(2), 165-171.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* East Norwalk, CT: Appleton-Century-Crofts.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience, 2*(10), 861-863.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Rajapakse, J. C., Vaituzis, A., Liu, H., et al. (1999). Development of the human corpus callosum during childhood and adolescence: A longitudinal MRI study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 23*(4), 571-588.

Gilhooly, K., & Logie, R. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods and Instrumentation, 12*(4), 395-427.

Gillberg, C., & Coleman, M. (1992). *The biology of the autistic syndromes* (2nd ed.). London: Mac Keith Press.

Gillott, A., Furniss, F., & Walter, A. (2001). Anxiety in high-functioning children with autism. *Autism, 5*(3), 277-286.

Goldberg, M. C., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: The effects of endogenous cueing and of distractors. *Developmental Science, 4*(2), 209-219.

Gollin, E. S. (1960). Developmental studies of visual recognition of incomplete objects. *Perceptual and Motor Skills, 11*, 289-298.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove, England: Lawrence Erlbaum.

Grandin, T. (2000). My experiences with visual thinking sensory problems and communication difficulties. from http://www.autism.org/temple/visual.html

Grier, J. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75*(6), 424-429.

Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. J. (1999). Executive functions in young children with autism. *Child Development, 70*(4), 817-832.

Gromko, J. E., & Poorman, A. S. (1998). The effect of music training on preschoolers' spatial-temporal task performance. *Journal of Research in Music Education, 46*(2), 173-181.

Guerin, F., Ska, B., & Belleville, S. (1999). Cognitive processing of drawing abilities. *Brain and Cognition, 40*(3), 464-478.

Gunter, H. L., Ghaziuddin, M., & Ellis, H. D. (2002). Asperger syndrome: Tests of right hemisphere functioning and interhemispheric communication. *Journal of Autism and Developmental Disorders, 32*(4), 263-281.

Gustafsson, L. (1997). Inadequate cortical feature maps: A neural circuit theory of autism. *Biological Psychiatry, 42*(12), 1138-1147.

Hadjikhani, N., Joseph, R. M., Snyder, J., Chabris, C. F., Clark, J., Steele, S., et al. (2004). Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. *Neuroimage, 22*(3), 1141-1150.

Hall, C. W., Gregory, G., Billinger, E., & Fisher, T. (1988). Field independence and simultaneous processing in preschool children. *Perceptual and Motor Skills, 66*(3), 891-897.

Hammill, D. D., Pearson, N., & Voress, J. (1993). *Developmental Test of Visual Perception* (2nd ed.). Austin, TX: PRO-ED.

Happé, F. (1994a). *Autism: An introduction to psychological theory*. London: UCL Press.

Happé, F. (1994b). Wechsler IQ profile and theory of mind in autism: A research note. *Journal of Child Psychology and Psychiatry, 35*(8), 1461-1471.

Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*(3), 843-855.

Happé, F. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions: A research note. *Journal of Child Psychology and Psychiatry, 37*(7), 873-877.

Happé, F. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology, 15*, 1-12.

Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences, 3*(6), 216-222.

Happé, F., & Booth, R. (2006). Sentence completion: a simple test of weak coherence in typical development, autism spectrum disorders and attention deficit/hyperactivity disorder. In preparation.

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*, In press.

Happé, F., Briskman, J., & Frith, U. (2001). Exploring the cognitive phenotype of autism: Weak "central coherence" in parents and siblings of children with autism: I. Experimental tests. *Journal of Child Psychology and Psychiatry, 42*(3), 299-307.

Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., et al. (1996). 'Theory of mind' in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport, 20*(9), 197-201.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, in press.

Hardy, R. C., Eliot, J., & Burlingame, K. (1987). Stability over age and sex of children's responses to Embedded Figures Test. *Perceptual and Motor Skills, 64*(2), 399-406.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(1), 163-169.

Haznedar, M., Buchsbaum, M. S., Wei, T.-C., Hof, P. R., Cartwright, C., Bienstock, C. A., et al. (2000). Limbic circuitry in patients with autism spectrum disorders studied with positron emission tomography and magnetic resonance imaging. *American Journal of Psychiatry, 157*(12), 1994-2001.

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry, 44*(4), 543-551.

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders, 35*(6), 787-793.

Heaton, P. (2006). Beyond perception: Musical representation and on-line processing in autism. Manuscript submitted for publication.

Heaton, P., Hermelin, B., & Pring, L. (1998). Autism and pitch processing: A precursor for savant musical ability. *Music Perception, 15*(3), 291-305.

Heaton, P., Pring, L., & Hermelin, B. (1999). A pseudo-savant: A case of exceptional musical splinter skills. *Neurocase, 5*(6), 503-509.

Heaton, P., & Wallace, G. L. (2004). Annotation: The savant syndrome. *Journal of Child Psychology and Psychiatry, 45*(5), 899-911.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory.* Oxford, UK: Wiley.

Heinze, H., Hinrichs, H., Scholz, M., Burchert, W., & Mangun, G. (1998). Neural mechanisms of global and local processing: A combined PET and ERP study. *Journal of Cognitive Neuroscience, 10*(4), 485-498.

Hermelin, B. (2001). *Bright splinters of the mind: A personal story of research with autistic savants.* Philadelphia, PA: Jessica Kingsley.

Hermelin, B., & O'Connor, N. (1967). Remembering of words by psychotic and subnormal children. *British Journal of Psychology, 58*(3-4), 213-218.

Hermelin, B., & O'Connor, N. (1970). *Psychological experiments with autistic children.* Oxford, UK: Pergamon.

Hermelin, B., & O'Connor, N. (1990). Art and accuracy: The drawing ability of idiot-savants. *Journal of Child Psychology and Psychiatry, 31*(2), 217-228.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124*(1), 62-82.

Hesse, F. W., & Spies, K. (1996). Effects of negative mood on performance: Reduced capacity or changed processing strategy? *European Journal of Social Psychology, 26*(1), 163-168.

Hill, E. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*(2), 189-233.

Hill, E., & Frith, U. (2003). Understanding autism: insights from mind and brain. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 28*(358), 281-289.

Hobson, R., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British Journal of Psychology, 79*(4), 441-453.

Holaday, M., Smith, D. A., & Sherry, A. (2000). Sentence completion tests: A review of the literature and results of a survey of members of the Society for Personality Assessment. *Journal of Personality Assessment, 74*(3), 371-383.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127*(4), 398-415.

Honda, H., Shimizu, Y., & Rutter, M. (2005). No effect of MMR withdrawal on the incidence of autism: a total population study. *Journal of Child Psychology and Psychiatry, 46*(6), 572-579.

Hooper, H. E. (1983). *The Hooper Visual Organization Test.* Los Angeles: Western Psychological Services.

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury Press.

Howlin, P., Baron-Cohen, S., & Hadwin, J. (1998). *Teaching children with autism to mind-read: A practical guide for teachers and parents.* New York: John Wiley & Sons.

Hoy, J. A., Hatton, C., & Hare, D. (2004). Weak central coherence: A cross-domain phenomenon specific to autism? *Autism, 8*(3), 267-281.

Hughes, C., Leboyer, M., & Bouvard, M. (1997). Executive function in parents of children with autism. *Psychological Medicine, 27*(1), 209-220.

Huss, E. T., & Kayson, W. A. (1985). Effects of age and sex on speed of finding embedded figures. *Perceptual and Motor Skills, 61*(2), 591-594.

Jacobsen, P. (2005). *Understanding how Asperger children and adolescents think and learn: creating manageable environments for AS students.* London: Jessica Kingsley.

Jarrold, C., Butler, D. W., Cottington, E. M., & Jimenez, F. (2000). Linking theory of mind and central coherence bias in autism and in the general population. *Developmental Psychology, 36*(1), 126-138.

Jarrold, C., Gilchrist, I. D., & Bender, A. (2005). Embedded figures detection in autism and typical development: Preliminary evidence of a double dissociation in relationships with visual search. *Developmental Science, 8*(4), 344-351.

Jarrold, C., & Russell, J. (1997). Counting abilities in autism: Possible implications for central coherence theory. *Journal of Autism and Developmental Disorders, 27*(1), 25-37.

Jensen, A. R., & Reynolds, C. R. (1983). Sex differences on the WISC-R. *Personality and Individual Differences, 4*(2), 223-226.

Jeyakumar, S. L., Warriner, E. M., Raval, V. V., & Ahmad, S. A. (2004). Balancing the need for reliability and time efficiency: Short forms of the Wechsler Adult Intelligence Scale-III. *Educational and Psychological Measurement, 64*(1), 71-87.

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry, 38*(5), 527-534.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition, 71*(2), 149-185.

Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired? *Psychological Medicine, 30*(5), 1169-1187.

Jolliffe, T., & Baron-Cohen, S. (2001a). A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate fragments of an object? *Cognitive Neuropsychiatry, 6*(3), 193-216.

Jolliffe, T., & Baron-Cohen, S. (2001b). A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate objects in context? *Visual Cognition, 8*(1), 67-101.

Jones, R. S., & Torgesen, J. K. (1981). Analysis of behaviors involved in performance of the Block Design Subtest of the WISC-R. *Intelligence, 5*(4), 321-328.

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43*(6), 807-821.

Just, M. A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain, 127*(8), 1811-1821.

Kail, R. (1990). *The development of memory in children* (3rd ed.). New York: W. H. Freeman.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.

Kalyan-Masih, V. (1985). Cognitive performance and cognitive style. *International Journal of Behavioral Development, 8*(1), 39-54.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217-250.

Karapetsas, A. B., & Vlachos, F. M. (1997). Sex and handedness in development of visuomotor skills. *Perceptual and Motor Skills, 85*(1), 131-140.

Karp, S. A., & Konstadt, N. L. (1963). *Manual for the Children's Embedded Figures Test.* Oxford, UK: Cognitive Tests.

Kaufman, A. S., Ishikuma, T., & Kaufman-Packer, J. L. (1991). Amazingly short forms of the WAIS--R. *Journal of Psychoeducational Assessment, 9*(1), 4-15.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children: Interpretive manual.* Circle Pines, MN: American Guidance Service.

Kee, D. W. (1994). Developmental differences in associative memory: Strategy use, mental effort, and knowledge access interactions. In H. W. Reese (Ed.), *Advances in child development and behavior, Vol 25* (pp. 7-32). San Diego, CA: Academic Press.

Keller, T. A., & Cowan, N. (1994). Developmental increase in the duration of memory for tone pitch. *Developmental Psychology, 30*(6), 855-863.

Kim, J. A., Szatmari, P., Bryson, S. E., Streiner, D. L., & Wilson, F. J. (2000). The prevalence of anxiety and mood problems among children with autism and Asperger syndrome. *Autism, 4*(2), 117-132.

Kimchi, R. (1990). Children's perceptual organisation of hierarchical visual patterns. *European Journal of Cognitive Psychology, 2*(2), 133-149.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin, 112*(1), 24-38.

Kimchi, R., & Palmer, S. E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance, 8*(4), 521-535.

Kinchla, R., & Wolfe, J. (1979). The order of visual processing: "Top-down," "bottom-up," or "middle-out." *Perception and Psychophysics, 25*(3), 225-231.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16*(2), 287-308.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809-816.

Kobayashi, S. (1985). An updated bibliography of picture-memory studies. *Perceptual and Motor Skills, 61*(1), 91-122.

Koch, C., Abbey, L., & Schmidt, S. (1995). Reexamining the Snodgrass and Corwin 1988 picture identification norms. *Perceptual and Motor Skills, 81*(3), 763-769.

Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. D. (2003). Electric brain responses reveal gender differences in music processing. *Neuroreport: For Rapid Communication of Neuroscience Research, 14*(5), 709-713.

Koffka, K. (1935). *Principles of gestalt psychology.* Oxford, UK: Harcourt, Brace.

Kohs, S. C. (1923). *Intelligence measurement: A psychological and statistical study based upon the block-design tests.* Oxford, UK: Macmillan.

Kramer, J. H., Delis, D. C., Kaplan, E., O'Donnell, L., & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology, 11*(4), 577-584.

Kramer, J. H., Ellenberg, L., Leonard, J., & Share, L. J. (1996). Developmental sex differences in global-local perceptual bias. *Neuropsychology, 10*(3), 402-407.

Kramer, J. H., Kaplan, E., Share, L., & Huckeba, W. (1999). Configural errors on WISC--III block design. *Journal of the International Neuropsychological Society, 5*(6), 518-524.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*(1), 159-179.

Lamb, M. R., Robertson, L. C., & Knight, R. T. (1990). Component mechanisms underlying the processing of hierarchically organized patterns: Inferences from patients with unilateral cortical lesions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 471-483.

Lawson, J., Baron-Cohen, S., & Wheelwright, S. (2004). Empathising and systemising in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 34*(3), 301-310.

Le Couteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Holdgrafer, M., et al. (1989). Autism Diagnostic Interview: A standardized investigator-based instrument. *Journal of Autism and Developmental Disorders, 19*(3), 363-387.

Lee, J. A., & Pollack, R. H. (1978). The effects of age on perceptual problem-solving strategies. *Experimental Aging Research, 4*(1), 37-54.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception and Psychophysics, 56*(4), 414-423.

Levy, I., Wallace, G., Black, D., Gilotty, L., Gibbs, M. C., Lee, P., et al. (2005). *Social and spatial memory in autism spectrum disorders and their relation to everyday behavior.* Paper presented at the International Meeting for Autism Research, Boston, Massachusetts.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431-461.

Lincoln, A. J., Allen, M. H., & Kilman, A. (1995). The assessment and interpretation of intellectual abilities in people with autism. In E. Schopler & G. B. Mesibov (Eds.), *Learning and cognition in autism* (pp. 89-117). New York: Plenum Press.

Lockyer, L., & Rutter, M. (1970). A five- to fifteen-year follow-up study of infantile psychosis: IV. Patterns of cognitive ability. *British Journal of Social and Clinical Psychology, 9*(2), 152-163.

Lopez, B., & Leekam, S. R. (2003). Do children with autism fail to process information in context? *Journal of Child Psychology and Psychiatry, 44*(2), 285-300.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview--Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659-685.

Lorsbach, T. C., & Reimer, J. F. (1997). Developmental changes in the inhibition of previously relevant information. *Journal of Experimental Child Psychology, 64*(3), 317-342.

Loth, E., Gómez, J. C., & Happé, F. (2006). Weak coherence in autism spectrum disorders: Specificity, individual differences and the relation between cognitive biases in the visual-spatial and verbal-semantic domains. Manuscript submitted for publication.

Lotter, V. (1966). Epidemiology of autistic conditions in young children: 1. Prevalence. *Social Psychiatry and Psychiatric Epidemiology, 1*(3), 124-137.

Lowe, R. C. (1973). A developmental study of part-whole relations in visual perception. *Journal of Genetic Psychology, 123*(2), 231-240.

Ludwig, T. E. (1982). Age differences in mental synthesis. *Journal of Gerontology, 37*(2), 182-189.

Lynn, R., & Gault, A. (1986). The relation of musical ability to general intelligence and the major primaries. *Research in Education, 36*, 59-64.

Lynn, R., & Mulhern, G. (1991). A comparison of sex differences on the Scottish and American standardisation samples of the WISC--R. *Personality and Individual Differences, 12*(11), 1179-1182.

Lynn, R., Wilson, R., & Gault, A. (1989). Simple musical tests as measures of Spearman's g. *Personality and Individual Differences, 10*(1), 25-28.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford CA: Stanford University Press.

Maestrini, E., Paul, A., Monaco, A. P., & Bailey, A. (2000). Identifying autism susceptibility genes. *Neuron, 28*(1), 19-24.

Mahlios, M. C., & D'Angelo, K. (1983). Group Embedded Figures Test: Psychometric data on children. *Perceptual and Motor Skills, 56*(2), 423-426.

Mandler, J. M. (1978). A code in the node: The use of story schema in retrieval. *Discourse Processes, 1*, 14-35.

Mandler, J. M., & Robinson, C. A. (1978). Developmental changes in picture recognition. *Journal of Experimental Child Psychology, 26*(1), 122-136.

Manjaly, Z. M., Marshall, J. C., Stephan, K. E., Gurd, J. M., Zilles, K., & Fink, G. R. (2003). In search of the hidden: an fMRI study with implications for the study of patients with autism and with acquired brain injury. *Neuroimage, 19*(3), 674-683.

Mann, T. A., & Walker, P. (2003). Autism and a deficit in broadening the spread of visual attention. *Journal of Child Psychology and Psychiatry, 44*(2), 272-284.

Martin, M. (1979). Local and global processing: The role of sparsity. *Memory and Cognition, 7*(6), 476-484.

Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: sensitivity and discrimination. *Ear and Hearing, 3*(6), 301-308.

McClelland, J. L. (2000). The basis of hyperspecificity in autism: A preliminary suggestion based on properties of neural nets. *Journal of Autism and Developmental Disorders, 30*(5), 497-502.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86*(5), 889-918.

McGivern, R. F., Huston, J., Byrd, D., King, T., Siegle, G. J., & Reilly, J. (1997). Sex differences in visual recognition memory: Support for a sex-related difference in attention in adults and children. *Brain and Cognition, 34*(3), 323-336.

McGuinness, D., & McLaughlin, L. (1982). An investigation of sex differences in visual recognition and recall. *Journal of Mental Imagery, 6*(1), 203-212.

McKelvey, J. R., Lambert, R., Mottron, L., & Shevell, M. I. (1995). Right-hemisphere dysfunction in Asperger's syndrome. *Journal of Child Neurology, 10*(4), 310-314.

Meili-Dworetzki, G. (1956). The development of perception in the Rorschach. In B. Klopfer (Ed.), *Development in the Rorschach technique*. New York: Harcourt, Brace and World.

Metzger, R. L., & Antes, J. R. (1983). The nature of processing early in picture perception. *Psychological Research, 45*(3), 267-274.

Metzger, R. L., & Perlmutter, M. (1984). Specific and global processing by preschool children and college adults. *Bulletin of the Psychonomic Society, 22*(4), 333-336.

Milne, E. (2003). *Investigating the correlates of weak central coherence in autism*. Unpublished PhD Thesis, University College London, London.

Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., & Plaisted, K. C. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry, 43*(2), 255-263.

Minshew, N. J., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychology and Psychiatry, 42*(8), 1095-1101.

Minshew, N. J., Turner, C. A., & Goldstein, G. (2005). The application of short forms of the Wechsler Intelligence scales in adults and children with high functioning autism. *Journal of Autism and Developmental Disorders, 35*(1), 45-52.

Mondloch, C. J., Geldart, S., Maurer, D., & de Schonen, S. (2003). Developmental changes in the processing of hierarchical shapes continue into adolescence. *Journal of Experimental Child Psychology, 84*(1), 20-40.

Morgan, B., Maybery, M., & Durkin, K. (2003). Weak central coherence, poor joint attention, and low verbal ability: Independent deficits in early autism. *Developmental Psychology, 39*(4), 646-656.

Moses, P., Roe, K., Buxton, R. B., Wong, E. C., Frank, L. R., & Stiles, J. (2002). Functional MRI of global and local processing in children. *Neuroimage, 16*(2), 415-424.

Mottron, L., & Belleville, S. (1993). A study of perceptual analysis in a high-level autistic subject with exceptional graphic abilities. *Brain and Cognition, 23*(2), 279-309.

Mottron, L., Belleville, S., & Menard, E. (1999). Local bias in autistic subjects as evidenced by graphic tasks: Perceptual hierarchization or working memory deficit? *Journal of Child Psychology and Psychiatry, 40*(5), 743-755.

Mottron, L., & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 131-148). Mahwah, NJ: Lawrence Erlbaum.

Mottron, L., Burack, J. A., Iarocci, G., Belleville, S., & Enns, J. T. (2003). Locally oriented perception with intact global processing among adolescents with high-functioning autism: Evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry, 44*(6), 904-913.

Mottron, L., Burack, J. A., Stauder, J. E., & Robaey, P. (1999). Perceptual processing among high-functioning persons with autism. *Journal of Child Psychology and Psychiatry, 40*(2), 203-211.

Mottron, L., Peretz, I., & Menard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry, 41*(8), 1057-1065.

Muhle, R., Trentacoste, S. V., & Rapin, I. (2004). The genetics of autism. *Pediatrics, 113*(5), 472-486.

Muller, R.-A., Behen, M., Rothermel, R., Chugani, D., Muzik, O., Mangner, T., et al. (1999). Brain mapping of language and auditory perception in high-functioning autistic adults: A PET study. *Journal of Autism and Developmental Disorders, 29*(1), 19-31.

Muller, R.-A., Pierce, K., Ambrose, J. B., Allen, G., & Courchesne, E. (2001). Atypical patterns of cerebral motor activation in autism: A functional magnetic resonance study. *Biological Psychiatry, 49*(8), 665-676.

Murphy, G. L., & Shapiro, A. M. (1994). Forgetting of verbatim information in discourse. *Memory and Cognition, 22*(1), 85-94.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353-383.

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica, 114*(3), 273-309.

Navon, D., & Norman, J. (1983). Does global precedence really depend on visual angle? *Journal of Experimental Psychology: Human Perception and Performance, 9*(6), 955-965.

Nelson, D. J., & Barresi, A. L. (1989). Children's age-related intellectual strategies for dealing with musical and spatial analogical tasks. *Journal of Research in Music Education, 37*(2), 93-103.

Norbury, C. F., & Bishop, D. V. (2002). Inferential processing and story recall in children with communication problems: A comparison of specific language impairment, pragmatic language impairment and high-functioning autism. *International Journal of Language and Communication Disorders, 37*(3), 227-251.

Norris, J. A., & Hoffman, P. R. (2002). Phonemic awareness: A complex developmental process. *Topics in Language Disorders, 22*(2), 1-34.

Nurmi, E. L., Dowd, M., Tadevosyan-Leyfer, O., Haines, J. L., Folstein, S. E., & Sutcliffe, J. S. (2003). Exploratory subsetting of autism families based on savant skills improves evidence of genetic linkage to 15q11-q13. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(7), 856-863.

Nygaard, L. C., & Pisoni, D. B. (1995). Speech perception: New directions in research and theory. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 63-96). San Diego, CA: Academic Press.

O'Connor, N., & Hermelin, B. (1987). Visual and graphic abilities of the idiot savant artist. *Psychological Medicine, 17*(1), 79-90.

O'Connor, N., & Hermelin, B. (1988). Low intelligence and special abilities. *Journal of Child Psychology and Psychiatry, 29*(4), 391-396.

O'Loughlin, C., & Thagard, P. (2000). Autism and coherence: A computational model. *Mind and Language, 15*(4), 375-392.

Olsson, I., Steffenburg, S., & Gillberg, C. (1988). Epilepsy in autism and autistic-like conditions. A population-based study. *Archives of Neurology, 45*(6), 666-668.

O'Riordan, M., & Plaisted, K. C. (2001). Enhanced discrimination in autism. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 54A*(4), 961-979.

Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R., Klin, A., et al. (2004). Performance on Cambridge Neuropsychological Test Automated Battery subtests sensitive to frontal lobe function in people with Autistic Disorder: Evidence from the collaborative programs of excellence in autism network. *Journal of Autism and Developmental Disorders, 34*(2), 139-150.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to Theory of Mind. *Journal of Child Psychology and Psychiatry, 32*(7), 1081-1105.

Ozonoff, S., Rogers, S. J., Farnham, J. M., & Pennington, B. F. (1993). Can standard measures identify subclinical markers of autism? *Journal of Autism and Developmental Disorders, 23*(3), 429-441.

Ozonoff, S., Strayer, D. L., McMahon, W. M., & Filloux, F. (1994). Executive function abilities in autism and Tourette syndrome: An information processing approach. *Journal of Child Psychology and Psychiatry, 35*(6), 1015-1032.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory and Cognition, 3*(5), 519-526.

Pasto, L., & Burack, J. A. (1997). A developmental study of visual attention: Issues of filtering efficiency and focus. *Cognitive Development, 12*(4), 427-439.

Patterson, M. B., Mack, J. L., & Schnell, A. H. (1999). Performance of elderly and young normals on the Gollin Incomplete Pictures Test. *Perceptual and Motor Skills, 89*(2), 663-664.

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S., et al. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain, 119*(1), 143-157.

Pellicano, E. (2004). *Investigating central coherence at the visuospatial level in typical development and in autism spectrum disorder.* Unpublished PhD Thesis, The University of Western Australia.

Pellicano, E., Gibson, L., Maybery, M., Durkin, K., & Badcock, D. R. (2005). Abnormal global processing along the dorsal visual pathway in autism: A possible mechanism for weak visuospatial coherence? *Neuropsychologia, 43*(7), 1044-1053.

Pellicano, E., Maybery, M., & Durkin, K. (2005). Central coherence in typically developing preschoolers: Does it cohere and does it relate to mindreading and executive control? *Journal of Child Psychology and Psychiatry, 46*(5), 533-547.

Pennings, A. (1988). The development of strategies in embedded figures tasks. *International Journal of Psychology, 23*(1), 65-78.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*(1), 51-87.

Penrose, L., & Penrose, R. (1958). Impossible objects: A special type of visual illusion. *British Journal of Psychology, 49*, 31-33.

Pezdek, K. (1980). Life-span differences in semantic integration of pictures and sentences in memory. *Child Development, 51*(3), 720-729.

Piaget, J. (1952). *The origins of intelligence in childhood.* New York: International Universities Press.

Piaget, J. (1969). *The mechanisms of perception.* London: Routledge & Kegan Paul.

Pickles, A., Bolton, P., Macdonald, H., Bailey, A., Le Couteur, A., Sim, C.-H., et al. (1995). Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: A twin and family history study of autism. *American Journal of Human Genetics, 57*(3), 717-726.

Pierce, S., & Bartolucci, G. (1977). A syntactic investigation of verbal autistic, mentally retarded, and normal children. *Journal of Autism and Childhood Schizophrenia, 7*(2), 121-134.

Plaisted, K. C. (2001). Reduced generalization in autism: An alternative to weak central coherence. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research.* (pp. 149-169). Mahwah, NJ: Lawrence Erlbaum.

Plaisted, K. C., O'Riordan, M., & Baron-Cohen, S. (1998a). Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *Journal of Child Psychology and Psychiatry, 39*(5), 765-775.

Plaisted, K. C., O'Riordan, M., & Baron-Cohen, S. (1998b). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry, 39*(5), 777-783.

Plaisted, K. C., Saksida, L., Alcantara, J., & Weisblatt, E. (2003). Towards an understanding of the mechanisms of weak central coherence effects: experiments in visual configural learning and auditory perception. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 358*(1430), 375-386.

Plaisted, K. C., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry, 40*(5), 733-742.

Polster, M. R., & Rapcsak, S. Z. (1994). Hierarchical stimuli and hemispheric specialization: Two case studies. *Cortex, 30*(3), 487-497.

Porporino, M., Shore, D. I., Iarocci, G., & Burack, J. A. (2004). A developmental change in selective attention and global form perception. *International Journal of Behavioral Development, 28*(4), 358-364.

Poulton, R. G., & Moffitt, T. E. (1995). The Rey--Osterrieth Complex Figure Test: Norms for young adolescents and an examination of validity. *Archives of Clinical Neuropsychology, 10*(1), 47-56.

Pring, L., & Hermelin, B. (1993). Bottle, tulip and wineglass: Semantic and structural picture processing by savant artists. *Journal of Child Psychology and Psychiatry, 34*(8), 1365-1385.

Pring, L., Hermelin, B., & Heavey, L. (1995). Savants, segments, art and autism. *Journal of Child Psychology and Psychiatry, 36*(6), 1065-1076.

Prior, M., & Hoffmann, W. (1990). Brief report: Neuropsychological testing of autistic children through an exploration with frontal lobe tests. *Journal of Autism and Developmental Disorders, 20*(4), 581-590.

Profita, J. T., Bidder, G., Optiz, J. M., & Reynolds, J. F. (1988). Perfect pitch. *American Journal of Medical Genetics, 29*(4), 763-771.

Rajapakse, J. C., Giedd, J. N., Rumsey, J. M., Vaituzis, A. C., Hamburger, S. D., & Rapoport, J. L. (1996). Regional MRI measurements of the corpus callosum: a methodological and developmental study. *Brain and Development, 18*(5), 379-388.

Redcay, E., & Courchesne, E. (2005). When is the brain enlarged in autism? A meta-analysis of all brain size reports. *Biological Psychiatry, 58*(1), 1-9.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: Some foundational issues. *Learning and Individual Differences, 7*(2), 145-162.

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30*(2), 178-191.

Richardson, A. (1977). Verbalizer-visualizer: A cognitive style dimension. *Journal of Mental Imagery, 1*(1), 109-125.

Richardson, J. A., & Turner, T. E. (2000). Field dependence revisited I: Intelligence. *Educational Psychology, 20*(3), 255-270.

Ridderinkhof, K., & van der Molen, M. W. (1995). A psychophysiological analysis of developmental differences in the ability to resist interference. *Child Development, 66*(4), 1040-1056.

Rinehart, N. J., Bradshaw, J. L., Moss, S. A., Brereton, A. V., & Tonge, B. J. (2000). Atypical interference of local detail on global processing in high-functioning autism and Asperger's disorder. *Journal of Child Psychology and Psychiatry, 41*(6), 769-778.

Ring, H. A., Baron-Cohen, S., Wheelwright, S., Williams, S. C., Brammer, M., Andrew, C., et al. (1999). Cerebral correlates of preserved cognitive skills in autism: A functional MRI study of Embedded Figures Task performance. *Brain, 122*(7), 1305-1315.

Roberts, J. A., Rice, M. L., & Tager-Flusberg, H. (2004). Tense marking in children with autism. *Applied Psycholinguistics, 25*(3), 429-448.

Robertson, D. A., Gernsbacher, M. A., Guidotti, S. J., Robertson, R. R. W., Irwin, W., Mock, B. J., et al. (2000). Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science, 11*(3), 255-260.

Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology, 23*(2), 299-330.

Robertson, L. C., Lamb, M. R., & Knight, R. T. (1988). Effects of lesions of temporal-parietal junction on perceptual and attentional processing in humans. *Journal of Neuroscience, 8*(10), 3757-3769.

Robinson, J., & Wilson, J. (1973). The impossible colonnade and other variations of a well-known figure. *British Journal of Psychology, 64*(3), 363-365.

Rock, I., Gopnik, A., & Hall, S. (1994). Do young children reverse ambiguous figures? *Perception, 23*(6), 635-644.

Rodgers, J. (2000). Visual perception and Asperger syndrome: Central coherence deficit or hierarchization deficit? A pilot study. *Autism, 4*(3), 321-329.

Roediger, H. L., & McDermott, K. B. (1996). False perceptions of false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(3), 814-816.

Ropar, D., & Mitchell, P. (1999). Are individuals with autism and Asperger's syndrome susceptible to visual illusions? *Journal of Child Psychology and Psychiatry, 40*(8), 1283-1293.

Ropar, D., & Mitchell, P. (2001). Susceptibility to illusions and performance on visuospatial tasks in individuals with autism. *Journal of Child Psychology and Psychiatry, 42*(4), 539-549.

Rourke, B. P. (Ed.). (1995). *Syndrome of nonverbal learning disabilities: Neurodevelopmental manifestations.* New York: Guilford Press.

Roux, F., & Ceccaldi, M. (2001). Does aging affect the allocation of visual attention in global and local information processing? *Brain and Cognition, 46*(3), 383-396.

Rowntree, J. P. (1973). A longitudinal study of musical development. *Psychology of Music, 1*(1), 29-32.

Rozencwajg, P., & Corroyer, D. (2002). Strategy development in a block design task. *Intelligence, 30*(1), 1-25.

Rumsey, J. M. (1992). Neuropsychological studies of high-level autism. In E. Schopler & G. B. Mesibov (Eds.), *High-functioning individuals with autism* (pp. 41-64). New York: Plenum Press.

Rumsey, J. M., & Hamburger, S. D. (1990). Neuropsychological divergence of high-level autism and severe dyslexia. *Journal of Autism and Developmental Disorders, 20*, 155-168.

Rutter, M. (1978). Diagnosis and definition of childhood autism. *Journal of Autism and Developmental Disorders, 8*(2), 139-161.

Rutter, M. (1983). Cognitive deficits in the pathogenesis of autism. *Journal of Child Psychology and Psychiatry, 24*(4), 513-531.

Rutter, M. (1989). Pathways from childhood to adult life. *Journal of Child Psychology and Psychiatry, 30*(1), 23-51.

Rutter, M. (2005). Incidence of autism spectrum disorders: Changes over time and their meaning. *Acta Paediatrica, 94*(1), 2-15.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2*(9), 437-442.

Sacks, O. (1995). Prodigies. In *An Anthropologist on Mars* (pp. 179-232). London: Picador.

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37*(1), 74-85.

Sattler, J. M. (1992). *Assessment of children: WISC-III and WPPSI-R supplement.* San Diego, CA: Author.

Savin, H., & Bever, T. (1970). The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior, 9*(3), 295-302.

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119*(1), 5-24.

Schellenberg, E., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General, 131*(4), 511-537.

Schellenberg, E., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14*(3), 262-266.

Scheuffgen, K. (1998). *Domain-general and domain-specific deficits in autism and dyslexia.* Unpublished PhD thesis, University College London, London.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*(2), 109-149.

Schorr, D., Bower, G. H., & Kiernan, R. J. (1982). Stimulus variables in the block design task. *Journal of Consulting and Clinical Psychology, 50*(4), 479-487.

Schwartz, D. W., & Karp, S. A. (1967). Field dependence in a geriatric population. *Perceptual and Motor Skills, 24*(2), 495-504.

Selfe, L. (1977). *Nadia: A case of extraordinary drawing ability in an autistic child.* London: Academic Press.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry, 24*(4), 613-620.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry, 34*(8), 1351-1364.

Shao, Y., Raiford, K. L., Wolpert, C. M., Cope, H. A., Ravan, S. A., Ashley-Koch, A. A., et al. (2002). Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *American Journal of Human Genetics, 70*(4), 1058-1061.

Shimmon, K., & Lewis, C. (2001). *Can children with autism be trained on the Tower of London task?* Paper presented at the BPS Developmental and Educational Sections Conference, Worcester, UK.

Shulman, G. L., Sullivan, M. A., Gish, K., & Sakoda, W. J. (1986). The role of spatial-frequency channels in the perception of local and global structure. *Perception, 15*(3), 259-273.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*(1), 120-136.

Simpson, G. B., & Foster, M. R. (1986). Lexical ambiguity and children's word recognition. *Developmental Psychology, 22*(2), 147-154.

Simpson, G. B., & Kreuger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language, 30*(6), 627-643.

Simpson, G. B., Krueger, M. A., Kang, H., & Elofson, A. C. (1994). Sentence context and meaning frequency effects in children's processing of ambiguous words. *Journal of Research in Reading, 17*(1), 62-72.

Simpson, G. B., & Lorsbach, T. C. (1983). The development of automatic and conscious components of contextual facilitation. *Child Development, 54*(3), 760-772.

Slater, A., Mattock, A., Brown, E., & Bremner, J. (1991). Form perception at birth: Cohen and Younger (1984) revisited. *Journal of Experimental Child Psychology, 51*(3), 395-406.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music.* Oxford, UK: Clarendon Press.

Smalley, S. L. (1998). Autism and tuberous sclerosis. *Journal of Autism and Developmental Disorders, 28*(5), 407-414.

Snodgrass, J. G., & Corwin, J. (1988). Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. *Perceptual and Motor Skills, 67*(1), 3-36.

Snodgrass, J. G., Smith, B., Feenan, K., & Corwin, J. (1987). Fragmenting pictures on the Apple Macintosh computer for experimental and clinical applications. *Behavior Research Methods, Instruments and Computers, 19*(2), 270-274.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174-215.

Snowling, M., & Frith, U. (1986). Comprehension in "hyperlexic" readers. *Journal of Experimental Child Psychology, 42*(3), 392-415.

Solso, R. L., & King, J. F. (1976). Frequency and versatility of letters in the English language. *Behavior Research Methods and Instrumentation, 8*(3), 283-286.

Soraci, S. A., & Carlin, M. T. (1992). Stimulus organization and relational learning. In N. W. Bray (Ed.), *International review of research in mental retardation* (Vol. 18, pp. 29-53). New York: Academic Press.

Sparrow, S., Balla, D., & Cichetti, D. (1984). *Vineland Adaptive Behavior Scales (VABS).* Circle Pines, MN: American Guidance Services.

Spencer, J., O'Brien, J., Riggs, K., Braddick, O., Atkinson, J., & Wattam-Bell, J. (2000). Motion processing in autism: Evidence for a dorsal stream deficiency. *Neuroreport: For Rapid Communication of Neuroscience Research, 11*(12), 2765-2767.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.

St George, M., Kutas, M., Martinez, A., & Sereno, M. (1999). Semantic integration in reading: Engagement of the right hemisphere during discourse processing. *Brain, 122*(7), 1317-1325.

Stark, M. E., & Coslett, H. B. (1993). The "Navon effect" -Forest before trees?: It depends on how old you are and how many trees are in the forest. *Journal of Clinical and Experimental Neuropsychology, 15*, 46-47.

Steel, J., Gorman, R., & Flexman, J. E. (1984). Neuropsychiatric testing in an autistic mathematical idiot-savant: Evidence for nonverbal abstract capacity. *Journal of the American Academy of Child Psychiatry, 23*(6), 704-707.

Stiles, J., Delis, D. C., & Tada, W. L. (1991). Global-local processing in preschool children. *Child Development, 62*(6), 1258-1275.

Stiles-Davis, J., Janowsky, J., Engel, M., & Nass, R. D. (1988). Drawing ability in four young children with congenital unilateral brain lesions. *Neuropsychologia, 26*(3), 359-371.

Stoecker, J. J., Colombo, J., Frick, J. E., & Allen, J. R. (1998). Long- and short-looking infants' recognition of symmetrical and asymmetrical forms. *Journal of Experimental Child Psychology, 71*(1), 63-78.

Street, R. F. (1931). *A gestalt completion test.* New York: Columbia University Press.

Streiner, D. L. (1996). Maintaining standards: differences between the standard deviation and standard error, and when to use each. *Canadian Journal of Psychiatry, 41*(8), 498-502.

Stromland, K., Nordin, V., Miller, M., Akerstrom, B., & Gilberg, C. (1994). Autism in thalidomide embryopathy: A population study. *Developmental Medicine and Child Neurology, 36*(4), 351-356.

Sudhalter, V., Maranion, M., & Brooks, P. (1992). Expressive semantic deficit in the productive language of males with fragile X syndrome. *American Journal of Medical Genetics, 43*(1-2), 65-71.

Swettenham, J., Baron-Cohen, S., Gomez, J. C., & Walsh, S. (1996). What's inside someone's head? Conceiving of the mind as a camera helps children with autism acquire an alternative to a theory of mind. *Cognitive Neuropsychiatry, 1*(1), 73-88.

Szatmari, P., Merette, C., Bryson, S. E., Thivierge, J., Roy, M.-A., Cayer, M., et al. (2002). Quantifying dimensions in autism: A factor-analytic study. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(4), 467-474.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins College Publishers.

Tada, W. L., & Stiles, J. (1996). Developmental change in children's analysis of spatial patterns. *Developmental Psychology, 32*(5), 951-970.

Tada, W. L., & Stiles-Davis, J. (1989). Children's analysis of spatial patterns: An assessment of their "errors" in copying geometric forms. *Cognitive Development, 4*(2), 177-195.

Tager-Flusberg, H. (1991). Semantic processing in the free recall of autistic children: Further evidence for a cognitive deficit. *British Journal of Developmental Psychology, 9*(3), 417-430.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113*(2), 345-361.

Taylor, E. (2001). The scientific contribution of Michael Rutter: An appreciation. In E. Taylor & J. Green (Eds.), *Research and innovation on the road to modern child psychiatry: Classic papers by Professor Sir Michael Rutter* (Vol. 2). London: Gaskell.

Taylor, S. (1973). Musical development of children aged seven to eleven. *Psychology of Music, 1*(1), 44-49.

Tekman, H. G., & Bharucha, J. J. (1992). Time course of chord priming. *Perception and Psychophysics, 51*(1), 33-39.

Tellegen, A., & Briggs, P. F. (1967). Old wine in new skins: Grouping Wechsler subtests into new scales. *Journal of Consulting Psychology, 31*(5), 499-506.

Térouanne, E. (1980). On a class of "impossible" figures: A new language for a new analysis. *Journal of Mathematical Psychology, 22*(1), 24-47.

Teuber, H.-L., & Weinstein, S. (1956). Ability to discover hidden figures after cerebral lesions. *Archives of Neurology and Psychiatry, 76*(4), 369-379.

Teunisse, J.-P., Cools, A. R., van Spaendonck, K. P., Aerts, F. H., & Berger, H. J. (2001). Cognitive styles in high-functioning adolescents with autistic disorder. *Journal of Autism and Developmental Disorders, 31*(1), 55-66.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*(4), 885-913.

Tillmann, B., & Bigand, E. (2004). The relative importance of local and global structures in music perception. *The Journal of Aesthetics and Art Criticism, 62*(2), 211-222.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception, 16*(1), 99-117.

Titchener, E. B. (1909). *Lectures on the experimental psychology of the thought-processes.* Oxford, UK: Macmillan.

Trehub, S. E. (2003). Absolute and relative pitch processing in tone learning tasks. *Developmental Science, 6*(1), 44-45.

Trehub, S. E., Schellenberg, E. G., & Hill, D. S. (1997). The origins of music perception and cognition: A developmental perspective. In I. Deliege & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 103-128). Hove, England: Psychology Press.

Tuchman, R., & Rapin, I. (1997). Regression in pervasive developmental disorders: seizures and epileptiform electroencephalogram correlates. *Pediatrics, 99*(4), 560-566.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory and Cognition, 22*(1), 111-126.

Uhlhaas, P. J., Silverstein, S. M., Phillips, W. A., & Lovell, P. G. (2004). Evidence for impaired visual context processing in schizotypy with thought disorder. *Schizophrenia Research, 68*(2-3), 249-260.

Van Giffen, K., & Haith, M. M. (1984). Infant visual response to Gestalt geometric forms. *Infant Behavior and Development, 7*(3), 335-346.

Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry, 33*(3), 489-507.

Vernon, P. E. (1972). The distinctiveness of field independence. *Journal of Personality, 40*(3), 366-391.

Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 45*(1), 135-170.

Volkmar, F. R., & Nelson, D. S. (1990). Seizure disorders in autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(1), 127-129.

Voyer, D., Voyer, S., & Bryden, M. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250-270.

Waber, D. P., & Holmes, J. M. (1985). Assessing children's copy productions of the Rey-Osterrieth Complex Figure. *Journal of Clinical and Experimental Neuropsychology, 7*(3), 264-280.

Waber, D. P., & Holmes, J. M. (1986). Assessing children's memory productions of the Rey-Osterrieth Complex Figure. *Journal of Clinical and Experimental Neuropsychology, 8*(5), 563-580.

Waiter, G. D., Williams, J. H., Murray, A. D., Gilchrist, A., Perrett, D. I., & Whiten, A. (2005). Structural white matter deficits in high-functioning individuals with autistic spectrum disorder: a voxel-based investigation. *Neuroimage, 24*(2), 455-461.

Warrington, E. K., & James, M. (1991). *Visual Object and Space Perception Test.* Bury St. Edmunds, UK: Thames Valley Test Company.

Wechsler, D. (1987). *Wechsler Memory Scale* (Revised ed.). London: The Psychological Corporation.

Wechsler, D. (1992). *Wechsler Intelligence Scale for Children* (3rd ed.). London: The Psychological Corporation.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale* (3rd ed.). London: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale* (3rd ed.). London: The Psychological Corporation.

Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences, 35*(4), 863-875.

Werner, H. (1957). *Comparative psychology of mental development.* Oxford, UK: International Universities Press.

Werner, L. A. (1996). The development of auditory behavior (or what the anatomists and physiologists have to explain). *Ear and Hearing. Special Issue: Maturation of the Auditory System, 17*(5), 438-446.

Wertheimer, M. (1938). Über Gestalttheorie [an address before the Kant Society, Berlin, 17th December, 1924] (W. D. Ellis, Trans.). In W. D. Ellis (Ed.), *A source book of Gestalt psychology.* Oxford, UK: Harcourt, Brace.

Whiteside, J. A., Elkind, D., & Golbeck, S. L. (1976). Effects of exposure duration on part-whole perception in children. *Child Development, 47*(2), 498-501.

Williams, D. (1992). *Nobody Nowhere.* London: Doubleday.

Williams, D. L., Goldstein, G., & Minshew, N. J. (2005). Impaired memory for faces and social scenes in autism: Clinical implications of memory dysfunction. *Archives of Clinical Neuropsychology, 20*(1), 1-15.

Williams, P., & Tarr, M. J. (1997). Structural processing and implicit memory for possible and impossible figures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(6), 1344-1361.

Wing, H. (1968). *Tests of musical ability and appreciation.* Cambridge, U.K: University Press.

Wing, L. (1976). *Early childhood autism.* London: Pergamon Press.

Wing, L. (1981). Sex ratios in early childhood autism and related conditions. *Psychiatry Research, 5*, 129-137.

Wing, L. (1991). The relationship between Asperger's syndrome and Kanner's autism. In U. Frith (Ed.), *Autism and Asperger syndrome* (pp. 93-121). New York: Cambridge University Press.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*(1), 11-29.

Witkin, H. A., & Goodenough, D. R. (1981). Cognitive styles: essence and origins. Field dependence and field independence. *Psychological Issues, 51*, 1-141.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the Embedded Figures Test.* Palo Alto, CA: Consulting Psychologists Press.

World Health Organization. (1993). *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research.* Geneva, Switzerland: Author.

Wundt, W. (1902). *Outlines of Psychology* (Trans., 2nd ed.). Oxford, UK: Engelmann.

Wyatt, B. S., Conners, F. A., & Carr, M. D. (1998). The Snodgrass picture fragment completion test: Alternate-form reliability. *Behavior Research Methods, Instruments and Computers, 30*(2), 360-368.

Young, A. W., & Deregowski, J. B. (1981). Learning to see the impossible. *Perception, 10*(1), 91-105.

Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development, 57*(3), 803-815.

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience, 6*(7), 692-695.

Zwaigenbaum, L., Szatmari, P., Jones, M. B., Bryson, S. E., MacLean, J. E., Mahoney, W. J., et al. (2002). Pregnancy and birth complications in autism and liability to the broader autism phenotype. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(5), 572-579.

Appendix A. Published developmental studies directly relevant to central coherence in TD populations

Reference	Participants (N, age)	Tasks	Main findings
Infancy			
Fantz, 1961, 1963	8 x < 2 days, 10 x 2-5 days, 25 x 2-6 months	Comparing relative fixation times to patterned (schematic face, concentric circles, newsprint) and non-patterned (white, yellow, red) surfaces.	Preference for visual patterns shown across all ages; twice as much time was spent looking at patterned stimuli than non-patterned. Longest fixation times to schematic face; evidence configurations are perceived in newborns.
L. B. Cohen & Younger, 1984	16 x 6-weeks, 16 x 3-months	Habituation paradigm using a single angle as stimulus. Infants tested on sensitivity to change in stimulus (e.g., same angle different orientation, different angle same line segments)	3-month-olds encode angles regardless of the orientation of the line segments; 6-week-olds encode the orientation of separate line segments that make up an angle, but not the whole form of the angle.
Van Giffen & Haith, 1984	32 x 1-month, 32 x 3-month	Eye tracking procedure to measure fixation activity to global forms (square, circle) composed of local elements	3-month-olds direct their fixations to a discrepant local element within an array of elements making up a square or a circle; thus showing sensitivity to violations of 'good form'. 1-month-olds did not direct their attention to such breaks in continuity.

Younger & Cohen, 1986	Exp 1: 12 x 4-months, 12 x 7-months Exp 2: 24 x 4-months, 24 x 7-months, 24 x 10-months	Habituation paradigm using drawings of imaginary animals that differed on three attributes (body, tail, feet); fixation time measured to test stimuli that consisted of correlated, uncorrelated, or novel attributes to the habituated stimulus.	4-month-olds habituate to both correlated and uncorrelated features and dishabituate to totally novel stimulus; thus respond to feature-specific information. 10-month-olds generalise to correlated features, but dishabituate to both uncorrelated and novel stimuli; thus attend to the relations among attributes (i.e., the whole object). (No clear evidence to suggest that 7-month-olds habituated or dishabituated; possibly a transitional stage).
Ghim & Eimas, 1988	Exp 1: 19 x 3-months, 13 x 4-months Exp 2: 44 x 3-months, 20 x 4-months	Assessing novelty preference: infants familiarised with single hierarchical figure; familiar stimulus then paired with novel stimulus that differed by local or global form. Tested an interference paradigm (Exp 2) by adding a second source of novelty to comparison stimuli (e.g., novel global/local form when testing for local/global familiarity respectively)	Exp 1: Infants able to process and remember both local and global level information; no difference in discriminability between the local and global forms. Exp 2: introduction of second source of novelty attracted infants' attention and disrupted preference test for novelty. Interference effects greater from novel global forms than novel local forms; evidence for global precedence.
Slater et al., 1991	Exp 1: 24 x newborns (6 hours – 5 days) Exp 2: 12 x newborns (13 hours – 4 days) Exp 2: 8 x newborns (14 hours – 5 days)	Habituation paradigm using simple two-line angles, test stimuli differ in angle or orientation (Exp 1: replication of Cohen & Younger, 1984). Exp 2 & 3: angle changed orientation over familiarisation trials (Exp 3: controlled for overall size of stimulus)	Exp 1: newborns dishabituated to a change of orientation but not to a change in angle. Exp 2 & 3: after repeated exposure of angle at different orientations, a strong novelty preference was shown to different angle. Therefore, when presented with changing input during familiarisation, newborn infants appear to be able to extract invariant global properties.

Freeseman et al., 1993	56 x 4-months	Discrimination and generalisation tasks using hierarchical figures	Infants process information in a global-to-local sequence: discrimination of global properties required less familiarisation time than for discrimination of local information.
Bhatt et al., 1994	24 x 6-months	Delayed recognition paradigm: testing whether alterations of incidental information (at global or local level) disrupt memory retrieval	6-month-olds showed generalisation of learning in contexts that consisted of similar global configuration but different local components to original learning environment.
Stoecker et al., 1998	96 x 4-months	Comparing "look duration" for symmetrical and asymmetrical stimuli	Infants with a general tendency for "short looking" processed symmetrical stimuli faster than asymmetrical stimuli suggesting a global visual scanning strategy (i.e., global-to-local processing sequence); whereas infants with "long looking" tendency, showed no such advantage suggesting less efficient scanning strategy (i.e., local-to-global).
Frick et al., 2000	Exp 1: 57 x 3-months Exp 2: 37 x 3-months	Paired comparison discrimination task: Infants familiarised with hierarchical array (Exp 1: 30/60 sec; Exp 2: 20 sec), then tested for recognition (novelty preference) by pairing array with a stimulus (i) of same global form but novel local elements, or (ii) of same local elements but different global form	Infants with long- and short-looking tendencies were sensitive to both global and local properties following 60 sec of familiarisation. At 30 sec, long-looking infants were sensitive to global (but not local) visual properties (akin to global precedence effects). This effect occurred in short-looking infants but only after 20 sec of familiarisation.

Childhood and Adolescence			
Meili-Dworetzki, 1956	122 x 3 to 15 years, 16 x adults.	Asked to describe six ambiguous figures that consisted of both wholes and parts (e.g., a teapot made of two human figures)	Children 3- to 6-years tended to report seeing the whole figure only. A shift to perceiving parts occurred 6-7 years, after which both parts and wholes were generally perceived simultaneously. Both parts and wholes were perceived in 80% of adults.
Elkind et al., 1964	23 x 4.5 years, 44 x 6 years, 50 x 7 years, 50 x 8 years, 28 x 9 years.	Asked to describe pictures of objects that were formed of identifiable local parts (e.g., a person made of individual pieces of fruit).	4- to 6-year-olds tended describe the parts only (i.e., "apple", "pear"); not until 8-years, did children reliably report both the parts and the whole spontaneously. No effect of gender or IQ.
Elkind et al., 1970	40 x 6-years	Construct part-whole configurations from memory; describe objects made from another item (e.g., a hat made of newspaper); and perform verbal and object tests of logical addition.	Children with tendency to report only parts (Low Combination, LC) compared with children who tended to report wholes or both parts and whole (High Combination, HC; from Elkind et al., 1964): LC children performed lower than HC on all measures: constructing part-whole figures from memory, tendency to not use expression "made of" when describing made-up objects; & tests of logical addition.
Lowe, 1973	40 x 4-5 years, 40 x 6-7 years, 40 x 8-9 years	Participants trained to match by shape, independent of size, by manually adjusting a comparison. Task required matching a local shape (square/rectangle) that comprised a global shape (rectangle/square).	The global shape influenced perception of parts such that the parts became to look like whole shape (i.e., small squares perceived as rectangles, and vice versa); this effect was greatest for 4 & 5-year-olds.

Whiteside, Elkind, & Golbeck, 1976	128 x 3-12 years	Similar task to Elkind, et al., (1964), but compound figure stimuli exposed at four different durations: 0.1, 1, 5, 10 seconds	Part plus whole and integrated part-whole responses increased with age and longer stimuli exposures. Exposure duration affected older children (greater than 10 years) less than younger children.
Carey & Diamond, 1977	Exp1 & 2: 12 x 6 years, 12 x 8 years, 12 x 10 years	Exp 1: Recognition of upright versus inverted unfamiliar faces and houses. Exp 2: Face recognition task with manipulations of facial expression, clothing, hairstyle, glasses, etc.	The face inversion effect not apparent until age 10 years. Children aged 6- & 8-years appear to encode unfamiliar faces in piecemeal fashion, shift towards configural processing found at 10 years.
Enns & Girgus, 1985	20 x 6 years, 20 x 8 years, 20 x 10 years, 20 x 11 years, 20 x 23 years	Distance estimation task: asked to judge distances between elements in patterns illustrating the Gestalt grouping principles of proximity, similarity, closure, and good continuation	All participants were sensitive to Gestalt patterns: distances between elements in a perceptual group consistently judged as smaller than reality. The magnitude of the distortions decreased significantly with age; suggests that ability to disregard Gestalt groupings improves with age when the task requires selective attention to elements.
Kimchi, 1990	14 x 3-5 years, 12 x 6-7 years, 14 x 8-9 years	Similarity judgment tasks (triads) using hierarchical figures (e.g., triangle made up of smaller triangles or squares)	All children showed sensitivity to stimulus density (i.e., a bias towards global matches with increase in number of local elements or size of comparison stimulus). No age group effects. Pattern of results similar to adult findings. Gender effects not explored.

Stiles et al., 1991	Exp 1: 60 x 3-year, 60 x 4-years, 60 x adults Exp 2: 20 x 3-year, 20 x 4-years, 20 x adults	Orientation judgment task using hierarchical figures; participants asked to judge which way an equilateral triangle was pointing under different contextual conditions: i.e., global & local level pattern elements; Exp 1 focused on global-to-local influences; Exp 2 on local-to-global influences	Children and adults were influenced by global level and local level cues; biasing effects were relatively stronger for adults; any response bias favouring the global level was weaker in young children; suggests that the degree of pattern integration was less marked for children than for adults.
Reyna & Kiernan, 1994	Exp 1 & 2: 25 x 6 years, 25 x 9 years	Story memory task: assessing recognition of actual sentences (verbatim), sentences consistent (gist) and not consistent (distractors) with meaning. Exp 1 instructed to recognise verbatim, Exp 2 instructed to process gist.	Exp 1: 9-year-olds had better verbatim memory than 6-year-olds; misrecognition based on gist was about the same in both age groups. Memory for surface form and gist appear to be independent: gist-based errors were not lower in 9-year-olds than 6-year-olds despite their better recognition of surface form. Exp 2: both age groups able to identify gist, 9-year-olds better than 6-year-olds. Verbatim and gist memory showed more dependency when instructed to process for gist.
Brainerd & Gordon, 1994	Exp 1: 48 x 5 years, 48 x 8 years Exp 2: 60 x 5 years, 60 x 8 years	Recognition memory for numerical information within a context: assessing verbatim (actual numbers) versus global gist (*most, least*) and pairwise gist (*more, less*).	Significant age group x memory test interaction: 5-year-old's verbatim memory was better than gist memory (both global & pairwise); 8-year-old's verbatim memory was worse than gist memory. Recall of verbatim information did not aid gist recall, even when explicitly instructed to remember the gist (Exp 2). Support for "parallel retrieval hypothesis" and independence of gist and verbatim recall.

Akshoomoff & Stiles, 1995a, 1995b	20 x 6 years, 20 x 7 years, 20 x 8 years, 20 x 9 years, 20 x 10 years, 20 x 12 years, adults (19-28 years)	Assessing drawing process and accuracy of copy and recall of the Rey Complex Figure	Drawing accuracy improved with age while drawing process changed with age. Between 6-9 years children tended to break the figure up into simple components; increased integration of the spatial array was shown with age. No significant age effects in starting strategy, although a trend for adults to start with perimeter more than younger children. Planning ability increased with age, with greater use of the large organisational units. Process used to copy figure influenced immediate recall: those who broke figure up into components tended to recall in a similar fashion and recall fewer elements. All participants recalled figure using more continuous lines than used in copy.
Dukette & Stiles, 1996	Exp 1: 28 x adults Exp 2 & 3: 24 x 4 years, 24 x 6 years, 24 x adults	Similarity judgment tasks (triads) using hierarchical figures: choices biased towards (i) global or (ii) local level, or both levels equally (iii) by identity and size, or (iv) either identity or size. Figures composed of letters (Exp 1 & 2) and geometric forms (Exp 3).	Children and adults both show a global level response bias with no difference between age groups. In local biasing condition all age groups show systematic shift towards local matching. In sparsity condition (i.e., lower density of local elements), 4-year-olds shifted to a local response bias, whereas 6-year-olds and adults showed same global response bias as standard condition. With geometric forms (Exp 3), both 4- & 6-year-olds showed this shift towards local matching in sparsity condition. Females made more global level matches than males in Exp 1 & 2, but not Exp 3.

Kramer et al., 1996	25 x 4-6 years, 28 x 7-8 years, 26 x 9-12 years	Similarity judgment tasks (triads) using hierarchical figures: comparison figures matched by local or global element (square/triangle); number and size of local elements varied	Two oldest age groups selected global match significantly more often than youngest group. Two oldest age groups also responded more globally as the number of local elements increased; youngest age group did not show this pattern. Males had a greater tendency towards the global form than females at all age groups.
Tada & Stiles, 1996	Exp 1, 2, & 3: 16 x 3-3.5 years, 16 x 3.5-4 years, 16 x 4-4.5 years, 16 x 4.5-5 years	Construction tasks: asked to draw whole form or complete a partially drawn figure; and to assemble each form using a limited or unlimited set of stimuli; Forms: Exp 1: + × ✱; Exp 2: cross-hatch grid (2 x 2 lines); Exp 3: large/small cross within a square	The number of parts produced for each form decreased with age; younger children treated line segments on each side of an intersection as separate parts and began to see them as a continuous unit with development. With difficult diagonal lines removed (Exp 2) youngest children still "over segmented" the form. Individual differences noted with some children attending to emergent feature of closure (Exp 3). Concluded that what constitutes a "part" changes with development. No effects of gender found.
Burack et al., 2000	20 x 6 years, 20 x 8 years, 20 x 10 years, 20 x 22 years	Visual search paradigm: detecting left/right position of targets defined by oblique orientation of four dots (2 pairs) amongst 1, 7, or 17 distractors (vertical column of four dots); spatial grouping of target pairs either short-range (local) or long-range (global)	Age improvements found for long-range targets, search for short-range targets constant across age. Pattern of performance the same across different levels of task difficulty (i.e. contrasting dot colour within and between pairs which affected access to local and global level information respectively).

Dukette & Stiles, 2001	Exp 1, 2 & 3: 12 x 4-years, 12 x 5-years, 12 x 6-years, 12 x 8-years, 12 x adults	Drawing hierarchical figures from memory and direct copy; density of local elements contrasted: standard (Exp 1), sparse (Exp 2), or dense (Exp 3); quality of reproduction of global and local level evaluated separately	Exp 1: 4-year-olds experienced difficulty reproducing global level information in memory condition compared to copy. Exp 2: all children (4 to 8 years) better at reproducing local than global information when density of stimuli was sparse (requiring integrative abilities). Exp 3: all children reproduced global and local levels equally well when number of local elements was dense. No gender effects found.
Mondloch et al., 2003	Exp 1: 24 x 6 years, 24 x 10 years, 24 x adults (18-28 years) Exp 2 & 3: 24 x 10 years, 24 x 14 years, 24 x adults (18-28 years)	Judged whether two briefly presented (Exp 1: 150 ms; Exp 2: 50 ms) hierarchical shapes where the same at the local or global level (attention directed to one level) Exp 3: Same task as Exp 2, with stimuli presented in the left or right hemisphere.	Global precedence effect present in all groups (faster reaction times to global level than local level) but was stronger in younger age groups, possibly reflecting an underdeveloped ability to ignore the irrelevant global level. 14-year-olds performed similarly to adults; stronger global precedence effect in 10-year-olds. Global processing was not strongly lateralised, but local processing was more efficient in the left hemisphere in all age groups (in a directed attention task).
Porporino et al., 2004	20 x 6 years, 20 x 8 years, 20 x 10 years, 20 x 12 years, 20 x adults (20-29 years)	Selective attention task using hierarchical figures; participants respond to square or diamond that could appear at local or global level; target figures flanked by distractors that were either neutral or in/congruent and in/compatible to target response.	Presence of distractors influenced global processing, but not local processing in younger children (6 to 8 years). Congruency and compatibility of distractor stimuli did not differentially affect local and global processing, or interact with age.

Adulthood/Aging			
Ludwig, 1982	Exp 1: 20 x 18-23 years, 20 x 65-70 years Exp 2: 24 x 17-21 years, 24 x 62-69 years Exp 3: 16 x 18-21 years, 16 x 61-70 years	Mentally combine two straight-line patterns and either reproduce (Exp 1) or recognise (Exp 2 & 3) the completed pattern. Requires global processing to mentally synthesize parts into whole.	Age decrement in ability to mentally synthesize: younger participants made more correct responses than older adults, even when matched on gender, handedness and memory ability (Exp 3). Older participants spent longer studying the individual patterns. No effect of gender.
Polster & Rapcsak, 1994	5 x 20-30 years, 5 x 60-69 years, 5 x 70-79 years	Selective attention task using briefly presented (100 ms) hierarchical figures: identifying whether H or S appeared at either local or global level. Attention directed to local or global level in different blocks of trials.	Young adults showed global advantage and global interference, participants in 60's showed no advantage and no interference, participants in 70's showed local advantage and local interference.
Roux & Ceccaldi, 2001	Exp 1: 20 x 19-25 years, 14 x 65-75 years Exp 2: 20 x 19-27 years, 17 x 60-80 years	Exp 1: Selective attention task using hierarchical figures: identifying whether H or S appeared at either local or global level. Exp 2: Divided attention task using hierarchical figures: identifying presence of H or S, which could appear equally at local or global levels, or be biased towards the local or global level within a block of trials.	Exp 1: Global advantage and global interference effects present in both age groups. Older participants more sensitive to interference caused by global information, and show a 'slowing down' in local identification, with a stronger global advantage. Exp 2: Global interference effects found in all three conditions (local bias, global bias, no bias) for both age groups. Manipulating the attentional bias had stronger effects on local/global advantage in younger group than older group.

Bruyer et al., 2003	Exp 1: 16 x 19-28 years, 16 x 67-77 years Exp: 2: 16 x 18-29 years, 16 x 60-79 years	Directed attention task replicating original Navon (1977) task; participants to decide whether "E/e" or "N/n" was displayed at a pre-specified level (global or local): Exp 1: Stimuli were global uppercase letters and local lowercase letters (or vice versa); Exp 2: upper & lower cases of global and local letters were randomly assigned.	Global precedence effects observed in both young and old participants in Exp 1 & Exp 2 where the random assignment of lower and upper case letters forced a late perceptual representation of the grapheme.

Appendix B. Pairwise matching between ASD and control participants

ASD group ID	Diagnosis	Age (years)	FIQ	VIQ	PIQ	Control group ID	Age (years)	FIQ	VIQ	PIQ	Age difference (years)	FIQ difference
301	Asperger	15.7	105	103	106	145	16.2	103	105	100	-0.54	2
302	Asperger	15.0	98	96	101	176	15.8	103	100	106	-0.79	-5
303	Autism	13.5	81	86	80	184	14.6	84	81	91	-1.13	-3
304	Asperger	14.7	104	104	103	7	13.9	107	95	118	0.79	-3
305	Asperger	13.7	98	103	92	119	13.8	95	105	85	-0.08	3
306	Asperger	13.5	134	132	124	190	14.1	122	132	106	-0.69	12
307	Asperger	12.9	89	88	94	9	12.1	88	84	97	0.76	1
308	Asperger	14.8	133	141	113	135	14.7	126	132	112	0.07	7
310	Asperger	16.2	108	108	106	144	16.3	105	103	106	-0.04	3
311	Asperger	15.9	81	89	76	174	15.3	86	84	91	0.61	-5
312	Asperger	14.8	100	111	88	117	14.0	93	105	83	0.75	7
313	Asperger	12.4	113	117	105	14	12.7	115	111	115	-0.32	-2
314	Asperger	16.0	94	107	81	141	16.3	98	111	86	-0.26	-4
315	Asperger	15.4	103	81	126	204	16.3	103	92	115	-0.85	0
316	Asperger	12.5	133	133	124	81	11.3	134	135	121	1.18	-1
317	Asperger	9.6	91	92	91	57	9.5	98	103	94	0.10	-7
318	Asperger	13.5	113	123	97	191	13.3	117	122	106	0.19	-4
319	Asperger	21.1	92	114	71	183	18.5	95	111	79	2.64	-3
320	Asperger	11.2	122	124	112	78	11.1	119	127	103	0.08	3

ASD group ID	Diagnosis	Age (years)	FIQ	VIQ	PIQ	Control group ID	Age (years)	FIQ	VIQ	PIQ	Age difference (years)	FIQ difference
321	Asperger	12.2	109	105	109	16	12.3	103	105	100	-0.07	6
322	Asperger	13.8	64	66	66	200	13.7	51	51	67	0.07	13
323	Asperger	14.5	52	62	54	199	15.8	47	60	48	-1.31	5
324	Asperger	13.8	78	84	79	185	13.0	75	87	70	0.77	4
325	Autism	14.2	63	70	65	175	15.0	73	73	80	-0.77	-10
326	Asperger	15.8	84	68	106	194	15.4	89	84	97	0.39	-5
327	Autism	15.5	95	103	88	125	14.4	92	105	80	1.10	3
329	Autism	18.0	82	81	88	207	19.1	89	84	97	-1.08	-7
330	Autism	18.0	63	73	61	212	20.3	73	81	71	-2.32	-10
331	Asperger	20.1	114	114	112	161	18.3	106	108	103	1.74	8
332	Asperger	16.6	54	68	51	198	16.0	60	51	80	0.53	-6
333	Autism	14.7	49	51	59	202	16.0	68	84	59	-1.28	-19

Appendix C. Task administration order

Session 1

[Questionnaire on Everyday examples of Central Coherence]
Drawing Task (copy of Rey/Pram figure)
Phoneme Segmentation Task
Block Design subtest from the WISC-III/WAIS-III
Homograph Reading Task
Information subtest from the WISC-III/WAIS-III
Drawing Task (recall of Rey/Pram figure)
[Immediate recall of story (Children's Memory Scale)]
Line Orientation Task
Sentence Completion Task
Un/segmented Block Design Task (whole designs)
[Kimchi Similarity Judgment Task]
[Delayed recall of story (Children's Memory Scale)]
[State Anxiety measure for Session 1]

Session 2

[Trait Anxiety and Depression measures]
Drawing Task (copy of Rey/Pram figure)
[Word Memory Unconnected List A]
[Click Migration Task]
[Same-Different Judgment Task: Geometric Figures]
Vocabulary subtest from the WISC-III/WAIS-III
Drawing Task (recall of Rey/Pram figure)
Un/segmented Block Design Task (segmented designs)
Pitch Identification Task
Embedded Figures Test
[Word Memory Connected List B]
[State Anxiety measure for Session 2]

Session 3

Picture Completion subtest from the WISC-III/WAIS-III
[Word Memory Unconnected List B]
Picture Memory Task: Immediate Recall and Recognition Test
Chord Segmentation Task
Control Task (to Pitch Identification and Line Orientation Tasks)
Fragmented Pictures Task
[Picture Memory Task: Delayed Recall]
Story Memory Task
[Same-Different Judgment Task: Real Figures]
[Word Memory Connected List A]
Navon Similarity Judgment Task
Homograph Reading Task: post-test
Impossible Figures Task
Chord Sequence Task
[State Anxiety measure for Session 3]

Note: Tasks in square brackets are not included in the present thesis

Appendix D. Language status of participants

(Age) group		*n*	English only	English first	English second
8-10	*male*	19	19	0	0
	female	35	31	4	0
11-13	*male*	21	15	3	3
	female	22	15	7	0
14-16	*male*	30	29	0	1
	female	21	20	0	1
17-25	*male*	27	24	2	1
	female	29	19	4	6
ASD		31	30	1	0
Controls		31	28	2	1

Appendix E. Music experience of participants

Typical Development

In order to measure the effects of musical training on task performance, information was gathered from each participant on formal musical training. Participants were questioned whether they had received individual music lessons (instrumental or singing), the number years they had received lessons and whether they had reached a specific grade or level of proficiency. Table E.1 presents the mean number of years of formal music training for male and female participants in each age group. Significant age group effects were found on this variable (χ^2 = 12.7, p = .005), with the youngest age group having received significantly fewer years of music experience than all older age groups (all z > 2.00, p < .04). There was no overall difference between male and female participants in the number of years of music training (z = 1.55, p = .12). Within age group analyses revealed that 14-16-year-old females had received significantly more years of musical training than male participants of this age (z = 2.54, p = .01). Between age group effects were not found for males (χ^2 = 3.63, p = .31), but were significant for females (χ^2 = 15.5, p = .001). Pairwise comparisons showed that females in the youngest age group had significantly fewer years of music experience than the three older female age groups (all z > 2.00, p < .05). Although the mean number of years of musical training appeared higher for females in the 14-16 year group than both the 11-13 and 17-25 year groups, neither difference reached conventional levels of significance (z = 1.90, p = .06, and z = 1.75, p = .08, respectively).

Table E.1 Number of years of formal musical training received by male and female TD participants in each age group and overall: Mean (*SD*).

Age group (years)		N	Years of music training
8-10	*male*	19	0.54 (0.65)
	female	35	0.73 (1.00)
11-13	*male*	21	1.00 (1.32)
	female	22	1.80 (1.90)
14-16	*male*	30	1.56 (1.85)
	female	21	3.67 (3.04)
17-25	*male*	27	2.52 (3.16)
	female	29	2.34 (3.01)
All	*male*	97	1.51 (2.17)
	female	107	1.97 (2.51)

There was a significant positive correlation between the number of years of music training and age in the TD sample (r_s = .22, p = .002; males r_s = .19, p = .06; females r_s = .26, p = .008). Significant positive correlations were also found between all IQ measures and the number of years of music training (r_s = .16 to .29, all p < .02; males r_s = .17 to .23, p < .10; females r_s = .20 to .40, p < .05).

Table E.2 presents participant characteristics in each age group split by music experience (more than one year versus less than one year of formal musical training). Musically trained participants were older ($t_{(202)}$ = 2.97, p = .003) and had significantly higher FIQ ($t_{(202)}$ = 3.37, p = .001) and VIQ ($t_{(202)}$ = 3.79, p < .001) than musically untrained participants. When split by age group, the effect of age was significant only in the youngest age group (z = 2.17, p = .03). This age effect is likely to be a consequence of having more opportunity to receive musical training; however musically trained participants also had significantly higher FIQ (z = 2.48, p = .01) and VIQ (z = 3.04, p = .002) than untrained participants within the 8-10 year age group. Music experience was associated with higher FIQ in oldest age group (z = 2.06, p = .04) and VIQ in 14-16 year group (z = 2.03, p = .04).

Table E.2 Participant characteristics of musically trained and untrained participants within each age group: Mean (*SD*).

Age group (years)	Music trained?	N	Age	FIQ	VIQ	PIQ
8-10	*Yes*	9	10.1 (0.3)	117.9 (13.2)	126.1 (14.9)	103.7 (11.9)
	No	45	9.6 (0.7)	105.1 (14.1)	107.5 (15.9)	100.9 (12.7)
11-13	*Yes*	18	12.5 (1.2)	109.4 (15.1)	110.5 (16.8)	105.0 (11.5)
	No	25	12.2 (0.8)	103.5 (15.7)	104.0 (15.6)	101.7 (15.6)
14-16	*Yes*	28	15.4 (0.8)	107.2 (12.1)	112.5 (13.9)	99.6 (10.9)
	No	23	15.6 (0.8)	101.5 (13.7)	102.6 (17.1)	100.0 (10.4)
17-25	*Yes*	25	20.6 (2.5)	115.8 (10.9)	115.3 (11.0)	112.3 (12.3)
	No	31	20.4 (2.3)	107.9 (15.6)	108.5 (15.3)	105.2 (17.1)
All	*Yes*	80	15.8 (4.0)	111.6 (13.1)	114.5 (14.4)	105.2 (12.5)
	No	124	13.9 (4.5)	104.8 (14.7)	106.1 (15.9)	102.0 (14.1)

ASD and Control groups

Table E.3 presents the mean number of years of music training for ASD and control participants, overall and split by high and low IQ. Control participants had received significantly more years of music training than ASD participants (z = 2.15, p = .03). This was also apparent when just including high-functioning participants (z = 2.32, p = .02), although low IQ groups had received an equivalent number of years of music training (z = 0.42, p = .67).

No significant correlation was found between age or MA and the number of years of music training in the ASD group (r_s = .15, p = .41; MA r_s = .08, p = .68) or the control group age (r_s = -.17, p = .36; MA r_s = .32, p = .09). Significant positive correlations were found between all IQ measures and years of music training in the control group (r_s = .38 to .49, all p < .05), but none were found in the ASD group (r_s = -.01 to .06, all p > .74).

Table E.3 Number of years of formal musical training received by ASD and Control participants overall and split by low and high IQ: Mean (*SD*).

	Group	*N*	Years of music training
All	ASD	31	0.56 (0.93)
	Control	31	1.44 (2.14)
High IQ	ASD	25	0.62 (0.96)
	Control	25	1.73 (2.29)
Low IQ	ASD	6	0.33 (0.82)
	Control	6	0.25 (0.42)

Table E.4 presents participant characteristics for ASD and control participants split by music experience. The ASD and control groups remained well matched on age and IQ after excluding participants who had received more than one year of formal musical training (all $t_{(48)} < .29$, $p > .78$). Musically trained control participants were matched in age to those musically trained in the ASD group ($z = 1.19$, $p = .23$), but were significantly higher in FIQ ($z = 1.98$, $p = .05$) and VIQ ($z = 1.93$, $p = .05$). Groups of musically trained and untrained participants were also compared within each group. In the ASD group, participants with and without music training were matched in age ($z = 1.24$, $p = .43$) and IQ (all $z < 1.08$, $p > .29$). In the control group, musically trained participants had higher IQ than untrained participants (all $z > 2.27$, $p < .03$), but were matched in age ($z = 1.28$, $p = .20$).

Table E.4 Participant characteristics of musically trained and untrained participants within ASD and Control groups: Mean (*SD*).

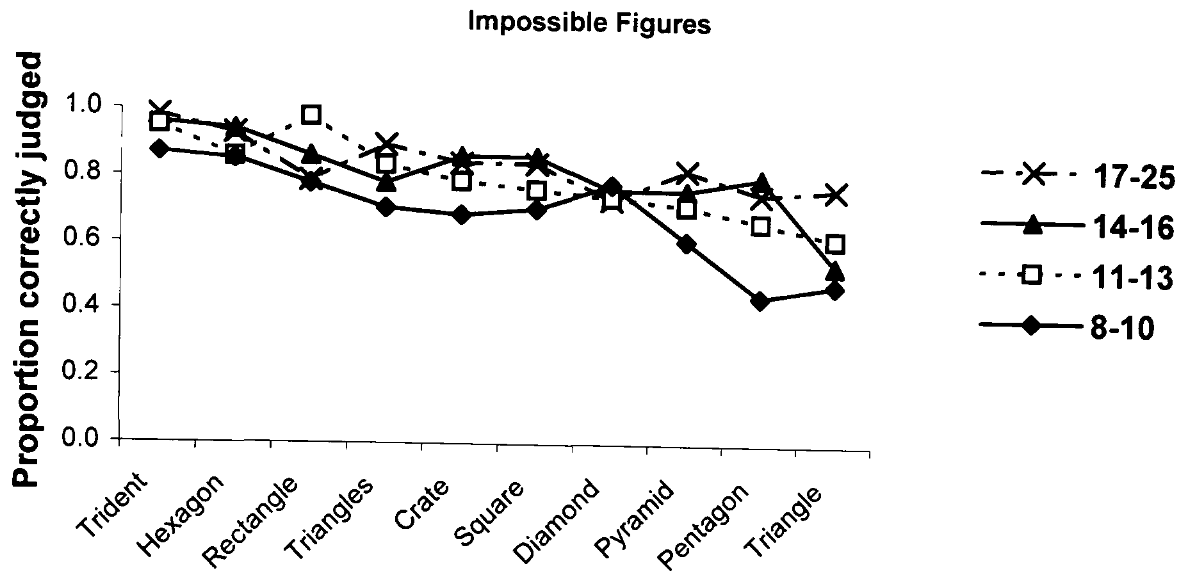
Group	Music trained?	N	Age	FIQ	VIQ	PIQ
ASD	*Yes*	5	15.7 (2.1)	85.6 (21.7)	87.2 (22.0)	87.7 (22.3)
	No	26	14.7 (2.4)	95.0 (23.8)	98.5 (23.0)	91.9 (21.3)
Control	*Yes*	11	14.2 (2.1)	106.3 (11.3)	108.8 (14.0)	101.7 (11.3)
	No	20	15.1 (2.5)	87.4 (22.3)	90.7 (22.9)	87.3 (18.9)

Appendix F. Un/segmented Block Design Task: Pilot data

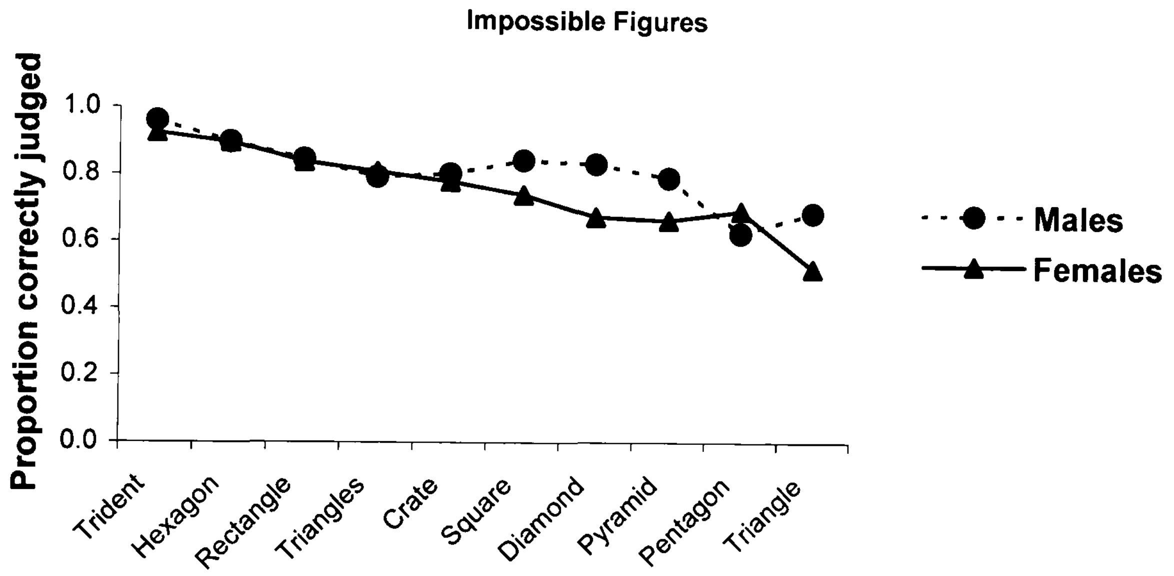
Pilot data showing accuracy and mean correct construction time for pilot data of 4-block and 9-block designs. Sets A and B, and C and D were equated on mean time for correct constructions.

	4-block designs	N	Proportion correct	Mean time for correct constructions (seconds)
SET A		67	0.97 (0.17)	14.6 (7.0)
		66	0.79 (0.41)	23.5 (16.3)
		77	0.96 (0.19)	21.8 (18.4)
		71	0.83 (0.38)	31.4 (20.4)
		Mean	**0.89 (0.29)**	**22.8 (15.5)**
SET B		9	1.00 (0.00)	16.4 (7.0)
		18	1.00 (0.00)	21.1 (23.9)
		21	1.00 (0.00)	26.8 (30.1)
		15	0.67 (0.49)	24.6 (10.0)
		Mean	**0.92 (0.12)**	**22.2 (17.7)**
	9-block designs	**N**	Proportion correct	Mean time for correct constructions (seconds)
SET C		13	0.77 (0.44)	66.8 (43.8)
		13	0.62 (0.51)	89.8 (48.1)
	Mean		**0.69 (0.48)**	**78.4 (46.0)**
SET D		15	0.60 (0.51)	65.7 (47.1)
		12	0.58 (0.51)	87.6 (36.1)
	Mean		**0.59 (0.51)**	**76.7 (41.6)**

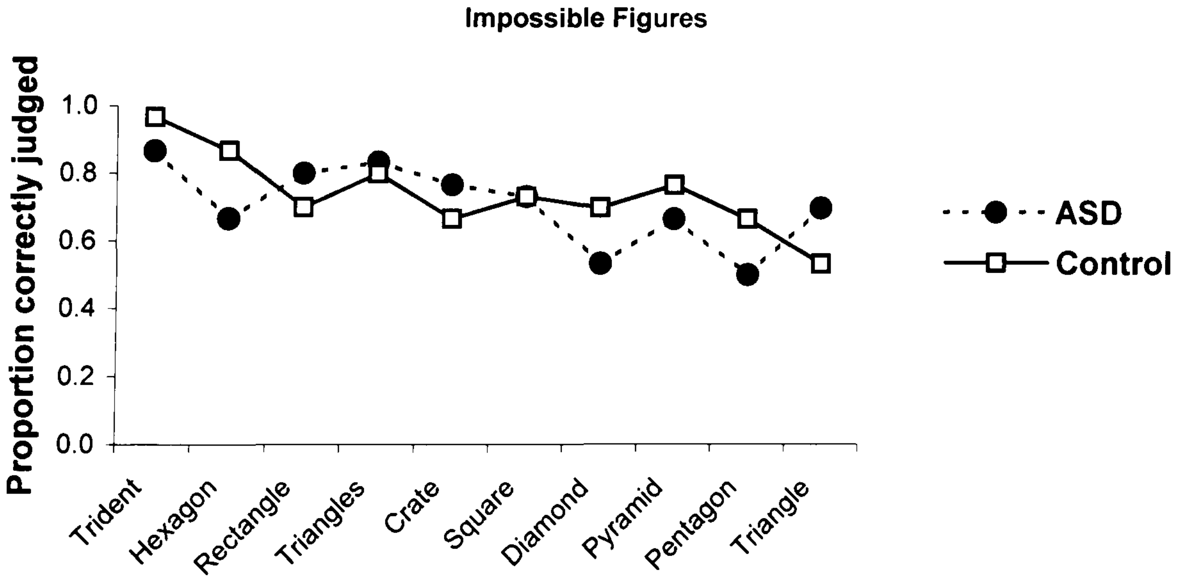
Appendix G. Impossible Figures Task: Item analysis



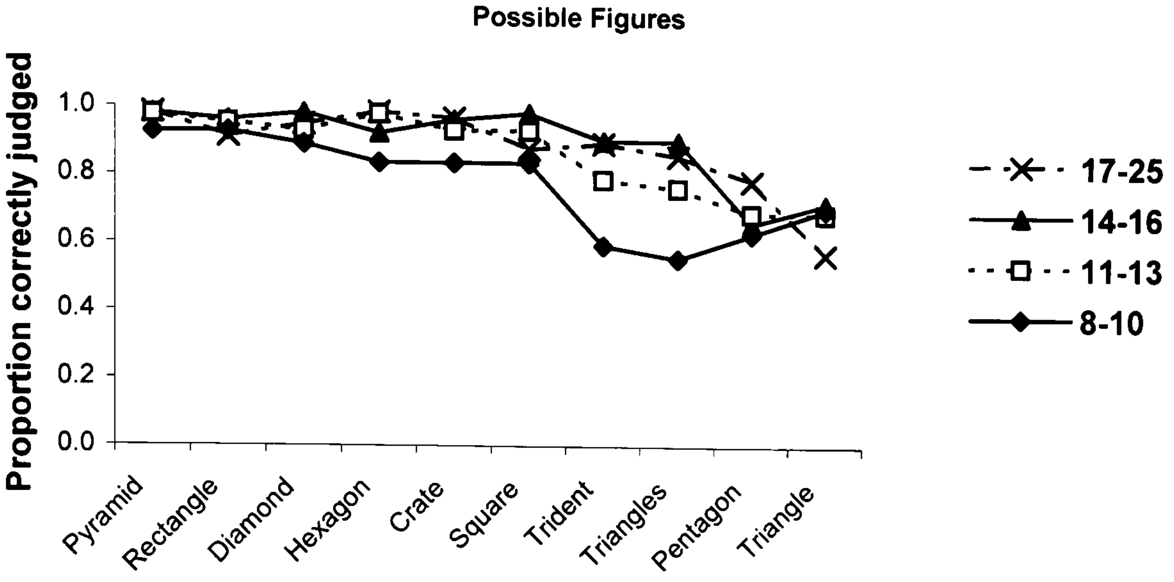
Rectangle: 8-10, 14-16, 17-25 < 11-13; Pentagon: 8-10 < 11-13, 14-16, 17-25; Triangle: 8-10, 14-16 < 17-25 (all $p < .05$).



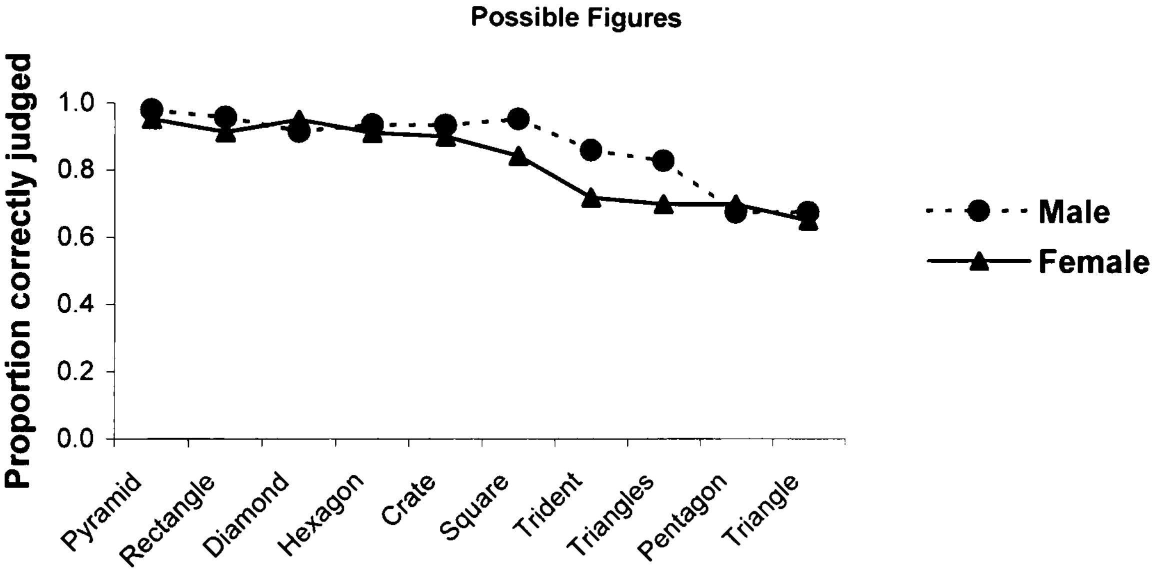
Females < Males: Diamond ($p = .009$) Pyramid ($p = .04$) Triangle ($p = .02$)



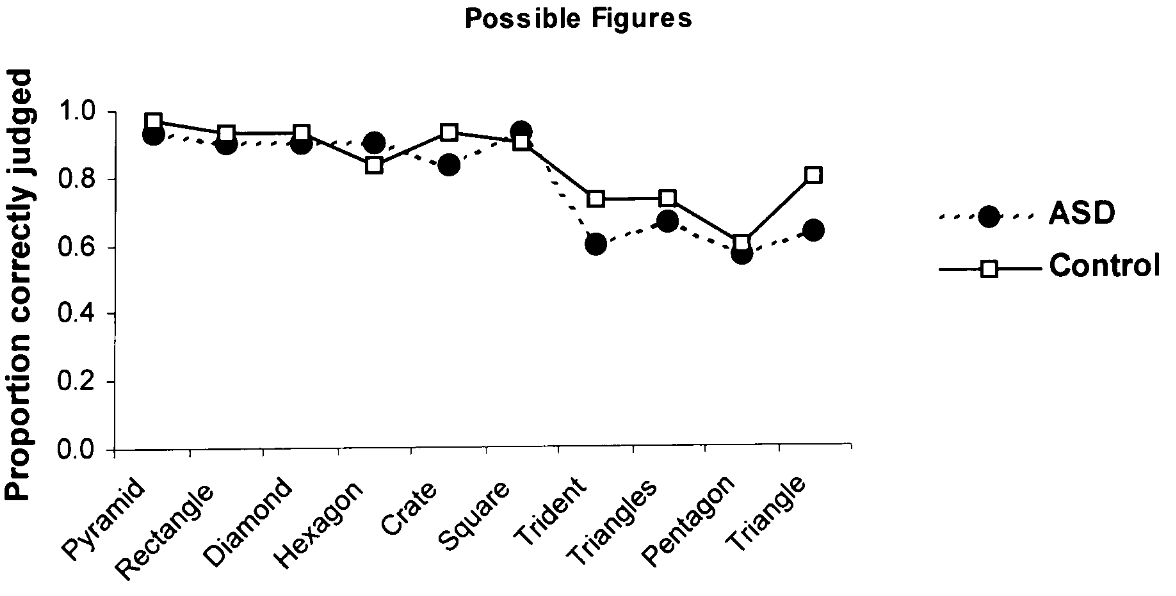
Hexagon: ASD < Control ($p = .07$)



Hexagon: 8-10 < 11-13, 17-25; Crate: 8-10 < 14-16, 17-25; Trident: 8-10 < 11-13, 14-16, 17-25; Triangles: 8-10 < 11-13, 14-16, 17-25 (all $p < .05$).



Females < Males: Square ($p = .008$) Trident ($p = .01$) Triangles ($p = .03$)



All ASD = Controls

Appendix H. Phoneme Segmentation Task: Stimuli

Target present	1 syllable	2 syllable	3 syllable
Initial sound	Pib (0) Poung (1) Prine (0) Plomb (0) Prite (0)	Payod (0) Putiff (0) Plebure (0) Protor(0) Piltin (0)	Pacoble (0) Posidise (5) Pluefernant (0) Pirresion (1) Prositye (2)
Medial	Lipt (0) Spow (2)	Freaspun (3) Upim (0) Soplen (2)	Elopive (0)
Final	Drup (1) Wep (1) Feap (2)	Dinope (1) Rishap (1)	Caviscope (4) Garovip (5) Alledap (7) Zeonasp (9)
Target absent	**1 syllable**	**2 syllable**	**3 syllable**
	Besh (0) Bood (1) Sluke (2) Wheeg (0) Gusf (0)	Kosbue (8) Migbus (0) Lindfy (2) Mufig (1) Subton (7)	Quaductric (7) Someder (0) Barano (2) Eledy (0) Mexonib (3)

Note: the number in brackets represents the number of pilot participants who incorrectly judged the target /p/ to be present or absent (maximum = 10).

Appendix I. Chord Segmentation Task: Stimuli

Trial Number	Chord	Comparison Note (degrees of scale)	Correct Response
Practice	C major	C (1st)	Yes
1	F# major	F# (1st)	Yes
2	G major	A (2nd)	No
3	B minor	B (1st)	Yes
4	G# minor	C# (4th)	No
5	E major	B (5th)	Yes
6	B♭ major	F♭ (5th)	Yes
7	E minor	F# (2nd)	No
8	F minor	F (1st)	Yes
9	D minor	G (4th)	No
10	A major	D (4th)	No
11	B♭ minor	C♭ (2nd)	No
12	G minor	D (5th)	Yes
13	C# minor	G# (5th)	Yes
14	E♭ major	A♭ (4th)	No
15	C# major	D# (2nd)	No
16	C major	C (1st)	Yes

Appendix J. Chord Sequence Task: Participant descriptives

Typical Development

The characteristics of TD participants who passed the inclusion criteria on the Chord Sequence Task are presented in Table J.1. Age groups were equivalent in FIQ and VIQ, but a significant effect of age group was found in PIQ ($F_{(3, 146)} = 4.29, p = .006$). The oldest age group had a significantly higher PIQ than the 14-16 year age group ($p = .006$). The oldest age group was also moderately higher in PIQ than the youngest age group ($p = .06$).

A significant effect of age group was found on the number of years of formal musical training (Kruskal Wallis test, $\chi^2 = 9.50, p = .02$). Subsequent Mann-Whitney U tests confirmed that the youngest age group had significantly fewer number of years of musical training than the two oldest age groups (all $z > 2.53, p < .02$), and moderately fewer than the 11-13 year group ($z = 1.81, p = .07$). The three older groups did not differ from one another in years of music experience.

Table J.1 Characteristics of participants who passed the inclusion criteria on the Chord Sequence Task by age group: Mean (*SD*).

Age group (years)	N	FIQ	VIQ	PIQ	Music training (years)
8 - 10	34	109.1 (15.2)	112.8 (17.5)	102.3 (11.7)	0.69 (0.80)
11 - 13	32	109.1 (15.6)	111.6 (14.6)	103.2 (15.0)	1.42 (1.64)
14 - 16	38	104.7 (12.3)	107.7 (15.2)	100.2 (10.4)	2.26 (2.39)
17 - 25	46	112.6 (14.3)	112.2 (14.1)	109.8 (14.7)	2.76 (3.17)

Participant characteristics divided by age group and gender, are presented in Table J.2. Chi-square tests showed that the distribution of male and female participants across the four age groups was equal ($\chi^2 = 3.78, p = .29$). Multivariate tests found no main effect of gender on age ($F_{(1, 142)} = 1.77, p = .19$). Male and female participants were also matched in age within each age group with no age group by gender interaction found ($F_{(3, 142)} = 2.34, p = .34$). A significant age group by gender interaction was observed for VIQ ($F_{(3, 142)} = 2.93, p = .04$). Independent samples t-tests showed a male advantage on VIQ within 17-25 year olds ($t_{(44)} = 2.27, p = .03$).

The mean number of years of music training was equivalent across all male and female participants who passed the inclusion criteria on the Chord Sequence Task ($z = 2.27, p = .07$). Within group analyses however showed that females in the 14-16 year age group had received significantly more years of music training than males in this age group ($z = 2.01, p = .05$).

Table J.2. Characteristics of participants who passed the inclusion criteria on the Chord Sequence task by age group and gender: Mean (*SD*).

Age group (years)		N	Age	FIQ	VIQ	PIQ	Music training (years)
8 - 10	*male*	13	9.7 (0.7)	113.5 (17.5)	119.6 (20.7)	102.8 (12.3)	0.73 (0.70)
	female	21	9.7 (0.6)	106.6 (13.1)	108.9 (13.9)	102.2 (11.3)	0.66 (0.85)
11 - 13	*male*	15	12.2 (0.9)	110.5 (16.3)	112.8 (13.6)	104.4 (16.7)	0.92 (1.01)
	female	17	12.4 (1.0)	107.8 (15.4)	110.5 (15.7)	102.1 (13.7)	1.86 (1.97)
14 - 16	*male*	23	15.6 (0.8)	103.2 (14.7)	104.7 (17.0)	100.8 (12.2)	1.57 (1.90)
	female	15	15.6 (0.9)	106.9 (7.3)	112.4 (11.1)	99.2 (7.2)	3.33 (2.72)
17 - 25	*male*	24	19.8 (2.1)	117.3 (15.1)	116.5 (14.8)	113.7 (14.6)	2.51 (3.24)
	female	22	20.8 (2.5)	107.4 (11.5)	107.5 (11.9)	105.5 (13.8)	3.05 (3.14)
All	*male*	75	15.2 (4.0)	110.9 (16.4)	112.7 (17.1)	106.0 (14.8)	1.59 (2.26)
	female	75	14.8 (4.7)	107.1 (12.1)	109.4 (13.2)	102.5 (12.1)	2.17 (2.53)

A significant effect of age group for the mean number of years of music training was found for females (χ^2 = 13.7, p = .003) but not for males. Females in the youngest age group had received significantly fewer number of years of music training than all older groups (all $z > 2.21, p < .03$), who did not differ significantly from one another.

ASD and Control groups

Participant descriptives for ASD and control individuals who passed inclusion criteria on the Chord Sequence Task is presented in Table J.3. The groups were equivalent in age and IQ (all $t_{(42)} < .83, p > .41$; high IQ $t_{(42)} < 1.32, p > .19$; low IQ $z < 1.73, p > .08$) and had received an equivalent number of years of formal music training (all $z = 1.55, p = .12$; high IQ $z = 1.86, p = .06$; low IQ $z = 0.19, p = .85$).

Table J.3. Characteristics of participants who passed the inclusion criteria on the Chord Sequence Task by group: Mean (*SD*).

	Group	N	Age	FIQ	VIQ	PIQ	Music training
All	ASD	22	15.3 (2.2)	93.9 (22.4)	97.2 (22.6)	91.5 (20.9)	0.70 (1.05)
	Control	22	14.7 (2.6)	95.4 (19.2)	99.1 (18.7)	92.6 (16.8)	1.57 (2.33)
High IQ	ASD	18	15.5 (2.3)	101.9 (15.6)	104.0 (18.9)	98.7 (15.0)	0.75 (1.08)
	Control	18	14.4 (2.5)	101.9 (13.7)	105.1 (13.7)	97.7 (13.5)	1.86 (2.48)
Low IQ	ASD	4	14.8 (1.2)	58.2 (6.3)	66.5 (3.4)	58.9 (7.9)	0.50 (1.00)
	Control	4	16.3 (2.9)	66.1 (10.5)	72.3 (14.7)	69.3 (8.5)	0.25 (0.50)

Appendix K. Pram and Rey Figure Scoring System

Table K.1. The 18 elements of the Rey figure.

Element	Description
1	Vertical Cross
2	Large Rectangle
3	Diagonal Cross
4	Horizontal Midline of Large Rectangle
5	Vertical Midline of Large Rectangle
6	Small Rectangle
7	Small Horizontal Line above Small Rectangle
8	Four Parallel Lines
9	Small Triangle above Large Rectangle
10	Small Vertical Line within Large Rectangle
11	Circle with Three Dots
12	Five Parallel Lines
13	Sides of Large Triangle attached to Large Rectangle
14	Diamond
15	Vertical Line within Sides of Large Triangle
16	Horizontal Line within Sides of Large Triangle
17	Horizontal Cross
18	Square attached to Large Rectangle

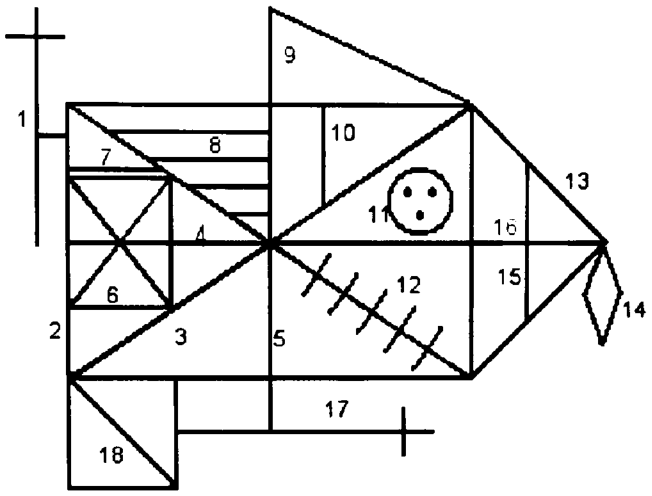


Table K.2. The four hierarchical levels of the Rey figure.

	Element	Description
Global external element	2	Large Rectangle
(score = 4)	13	Sides of Large Triangle attached to Large Rectangle
Global internal element	3	Diagonal Cross
(score = 3)	4	Horizontal Midline of Large Rectangle
	5	Vertical Midline of Large Rectangle
	16	Horizontal Line within Sides of Large Triangle
Local perimeter element	1	Vertical Cross
(score = 1)	9	Small Triangle above Large Rectangle
	14	Diamond
	17	Horizontal Cross
	18	Square attached to Large Rectangle
Local internal element	6	Small Rectangle
(score = 0)	7	Small Horizontal Line above Small Rectangle
	15	Vertical Line within Sides of Large Triangle
	8	Four Parallel Lines
	10	Small Vertical Line within Large Rectangle
	11	Circle with Three Dots
	12	Five Parallel Lines

Table K.3. The 16 elements of the Pram figure.

Element	Description
1	Trapezium
2	Vertical midline
3	Handle
4	Left diagonal line
5	Right diagonal line
6	Rim
7	Diamond
8	Crescent
9	Right arc
10	Two radius lines
11	Four dots
12	Three lines connecting four dots
13	Bottom horizontal line
14	Outer wheel rim
15	Inner wheel rim
16	Four spoke lines

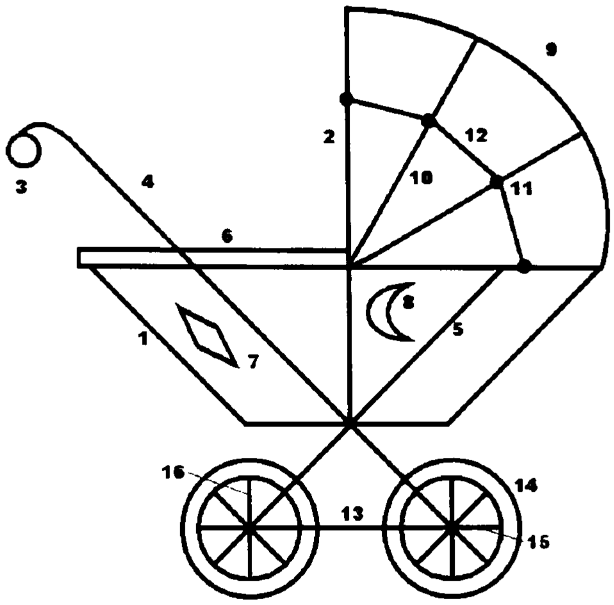


Table K.4. Description of the 18 elements of the Pram figure and the scoring criteria for accuracy.

Element	Description	Scoring Criteria
1	Trapezium	The base of the trapezium should be between ⅓ to ½ of its maximum length; the height of the trapezium should not be greater than its base; the overall shape should be symmetrical.
2	Vertical midline	Must be a continuous vertical line from the top of the pram and meet in the intersection of (4) and (5) at the base of the trapezium. the height of the top vertical line must be clearly longer than the height of the trapezium.
3	Handle	Must be circular in shape; connected to (4) and be positioned at approximately half the height of the top half of the pram and extend over the left of the Pram.
4	Left diagonal line	Must be a continuous straight line from the handle (3) to the centre of the right wheel, drawn through (6), intersecting (2) and (5), and be parallel to left side of trapezium (1).
5	Right diagonal line	Must be a continuous straight line starting from left of the bottom dot (11), to the centre of the left wheel, intersecting (2) and (4) and be parallel to the right side of trapezium (1).
6	Rim	The edge must extend slightly over the top left corner of the trapezium, and end at (2).
7	Diamond	Must be diamond in shape and not touch the sides of (1) or (4).
8	Crescent	Crescent shape must face to the right and not touch the sides of (1), (2) or (5).
9	Right arc	Must be a smooth curve and connect from the top right corner of (1) and the top of (2)
10	Two radius lines	Both must begin at the intersection of (1) and (2) and divide the quarter circle into approximately three equal parts.
11	Four dots	The four dots must be positioned approximately two-thirds along the radii of arc (9) and be of equal size.
12	Three lines connecting four dots	Three lines must be purposely straight, as compared to the curve of (9), and connect each dot.
13	Bottom horizontal line	Line must connect the centre point of each wheel and correspond (in length and position) to the base of trapezium (1).
14	Outer wheel rim	The outer wheel rim must be approximately circular and of similar size for both wheels; the wheels must not extend beyond the maximum length of the trapezium (1).

15	Inner wheel rim	The inner wheel rim must also be approximately circular and of similar size for both wheels.
16	Four spoke lines	The spoke lines must end at the inner wheel rim (15) intersect in the middle, and be of similar length; with two diagonal lines, one vertical and one horizontal.

Table K.5. The four hierarchical levels of the Pram figure.

	Element	**Description**
Global external element	1	Trapezium
(score = 4)	9	Right arc
Global internal element	2a	Vertical midline (top)
(score = 3)	2b	Vertical midline (middle)
	4a	Left diagonal line (top with handle)
	4b	Left diagonal line (middle)
	4c	Left diagonal line (bottom)
	5a	Right diagonal line (middle)
	5b	Right diagonal line (bottom)
	13	Bottom horizontal line
Local perimeter element	15	Inner wheel rim
(score = 1)	14	Outer wheel rim
	6	Rim
	3	Handle
Local internal element	10	Two radius lines
(score = 0)	7	Diamond
	8	Crescent
	11	Four dots
	12	Three lines connecting four dots
	16	Four spoke lines

Drawing Style: Inter-rater reliability

The inter-rater reliability of the scoring schedule devised for the Pram and Rey figures was investigated. A second-coder independently rated drawings from 48 participants, while being blind to age and group membership. Measures of style and order of construction were taken from the colour coding made by the researcher, while accuracy was rated from the original pictures. Thirty-eight participants were from the TD sample (19%), and comprised nine participants from each of the 11-13 and 17-25 year groups, and ten participants from each of the 8-10 and 14-16 year groups. An equivalent number of male and female participants were double-coded in each age group. Eight TD participants were also matched control participants to members of the ASD group. Drawings from a further two control participants were double-coded, resulting in inter-rater reliability data for 32% of the control group (10 from 31) and 26% of the ASD group (8 from 31). It was ensured that the selected participants were a representative sample from the low and high IQ groups.

Rey figure

Accuracy scores for copy and recall of the Rey figure showed high inter-rater agreement, as indicated by correlations ranging from .82 to .97 for participants in the TD, ASD, and control groups. Good agreement was also found on individual elements of the Rey figure with percentage agreement of 80% for both copy and recall items across all participants. The *order of construction* index for the Rey figure correlated significantly between the coders for copy and recall conditions in all groups (r_s ranged from .66 to .98). Cohen's Kappa coefficients indicated moderate to high inter-rater agreement for five of the six components scored for *style*, ranging from κ = .37 to .93 for copy, and κ = .47 to .83 for recall. The notable exception was the component incorporating the large triangle and vertical line within the large triangle (elements 13 and 15), which had low agreement for copy and recall (κ < .27). Further analysis revealed a systematic difference between the two coders for this component: As elements 13 and 15 are not typically drawn consecutively, the first-coder rated for line continuity only, whereas the second-coder took both dimensions into account resulting in consistently lower scores.

Pram figure

Accuracy scores for copy and recall of the Pram figure showed high inter-rater agreement as demonstrated by correlations ranging from .76 to .95 for groups of TD, ASD, and control participants. Good agreement was also demonstrated across all participants for the elements of the Pram figure, with 81% of copy items and 74% of recall items rated identically. The *order of construction* index also showed good inter-rater agreement with significant correlations between the two coders for copy and recall conditions (r_s = .61 to .80). Cohen's Kappa coefficients revealed moderate to high agreement between coders for the six components rated for *style* (κ = .43 to 1.00).

In sum, the level of agreement between the two coders was found to be adequate for both the Rey and Pram figures, with comparable performance across groups of TD, ASD, and control participants. In order to maintain consistency, ratings for all picture made by the first-coder (the author) were used in the analyses.

Appendix L. Picture Memory Task: Inter-rater reliability

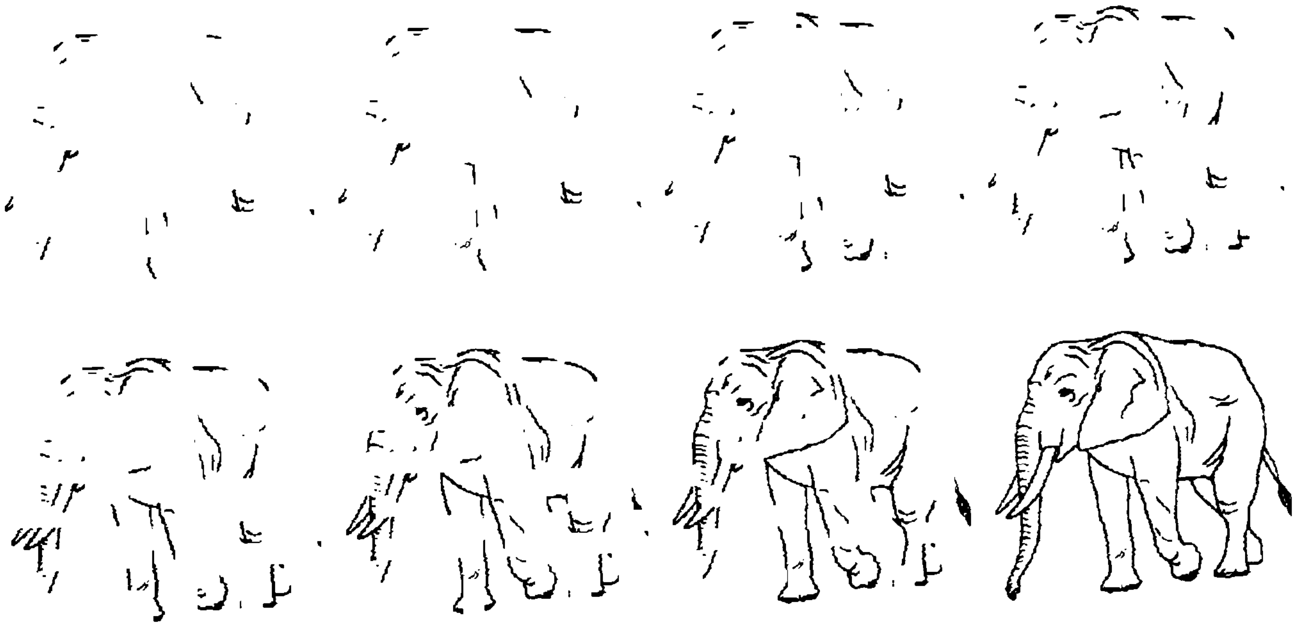
A second-coder, who was blind to group membership and hypotheses, independently rated descriptions from the Picture Memory Task from 30% of the TD sample (61 from 204), 30% of the ASD group (10 from 30), and 35% of the control group (11 from 31). Good inter-rater agreement was found for the tallies of local features. Correlation coefficients for the total number of objects listed, references to surface detail, and object position ranged from .82 to .97 in TD, .85 to .98 in ASD, and .73 to .99 in the control group. Cohen's Kappa coefficients indicated high inter-rater agreement for judging whether a global description was given initially or ever in the TD group ($\varkappa$ = .83 to .93) and in the ASD group ($\varkappa$ = .91 to 1.00). This index showed moderate agreement in the control group ($\varkappa$ = .57 to .78).

The frequency of inferences, personal responses, and stating objects not present was low across all groups and was combined across immediate and delayed recall for inter-rater analyses. The most subjective rating was for inference comments, which differed between raters for 15% of TD participants (9/61) and 30% of ASD participants (3/10). Ratings for inference comments did not differ for control participants. Ratings for the number of personal comments differed for 5% of TD participants (3/61), 20% of ASD participants (2/10) and no control participants. Finally, ratings for objects not present disagreed for 7% of TD participants (4/61), 10% of ASD participants (1/10), and 9% of control participants (1/11).

Appendix M. Fragmented Pictures Task: Stimuli

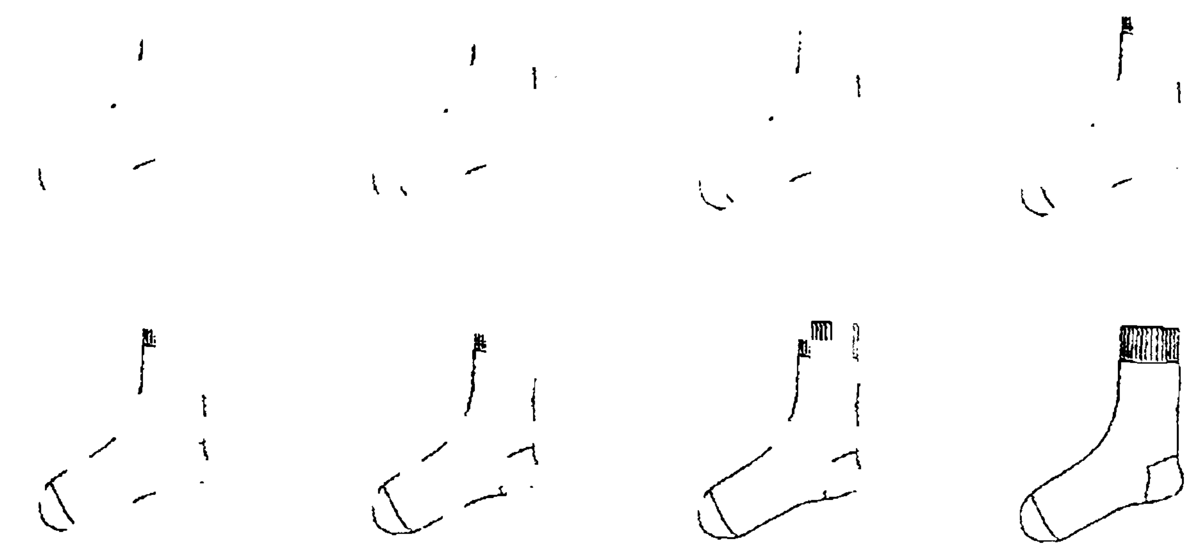
(1) Apple

(2) Elephant

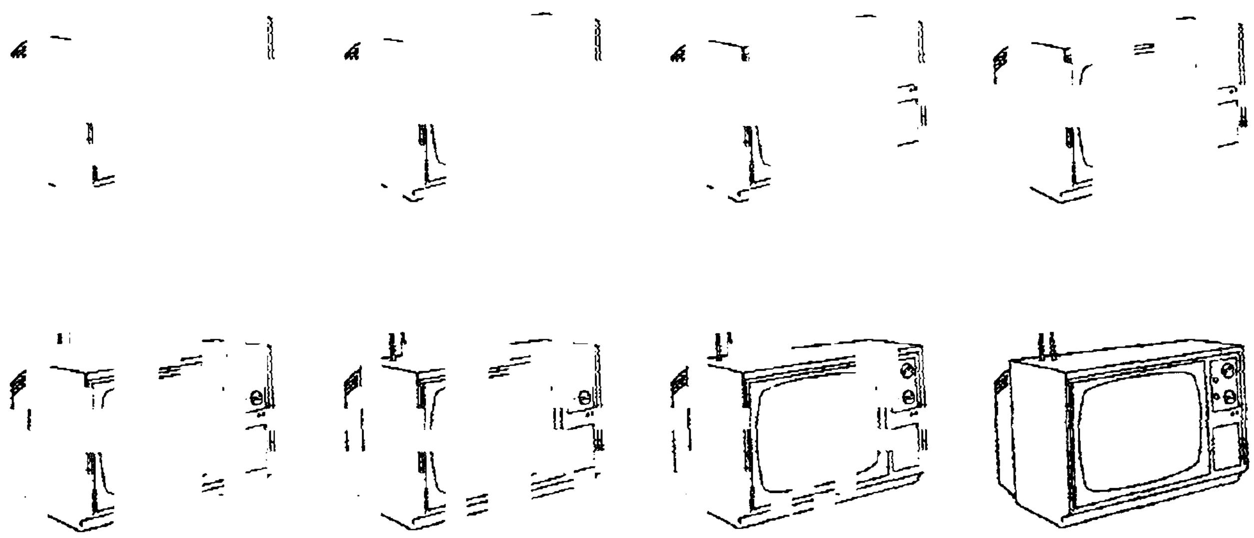


(3) Pig

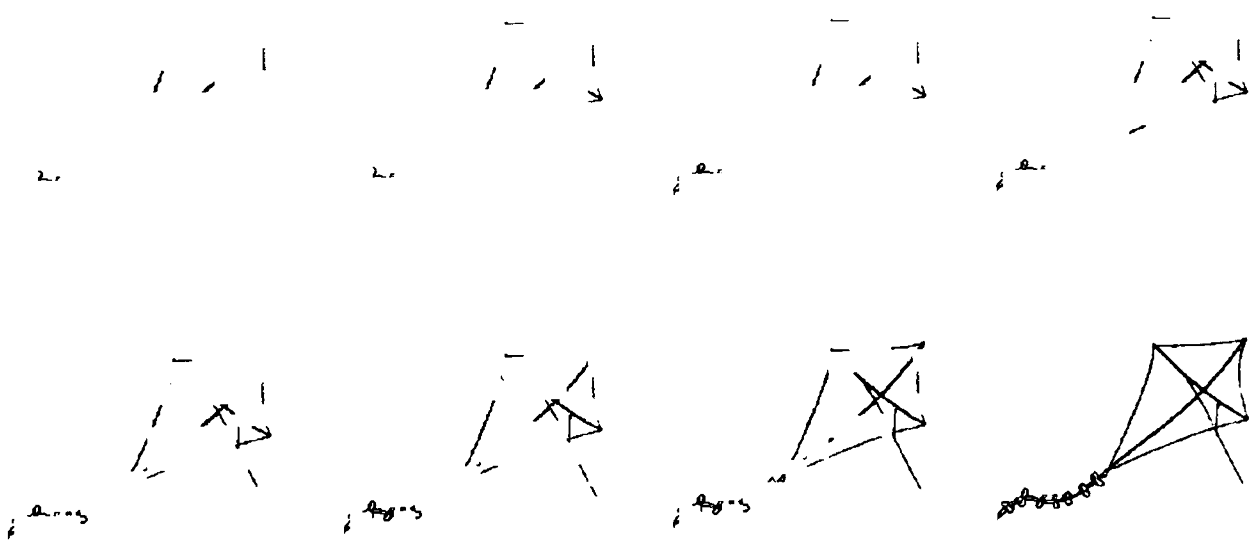
(4) Sock



(5) Television

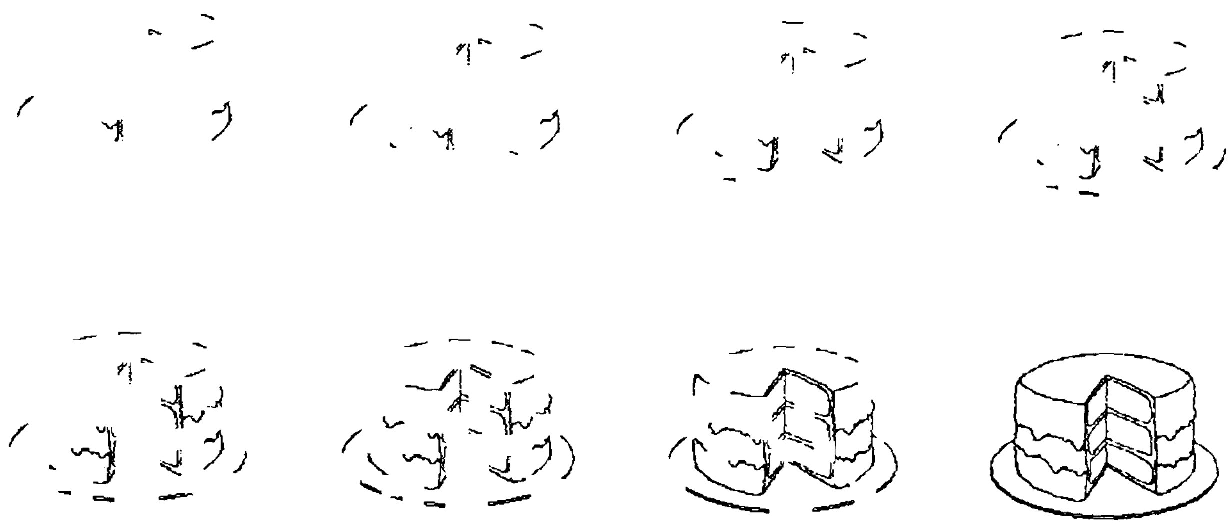


(6) Kite

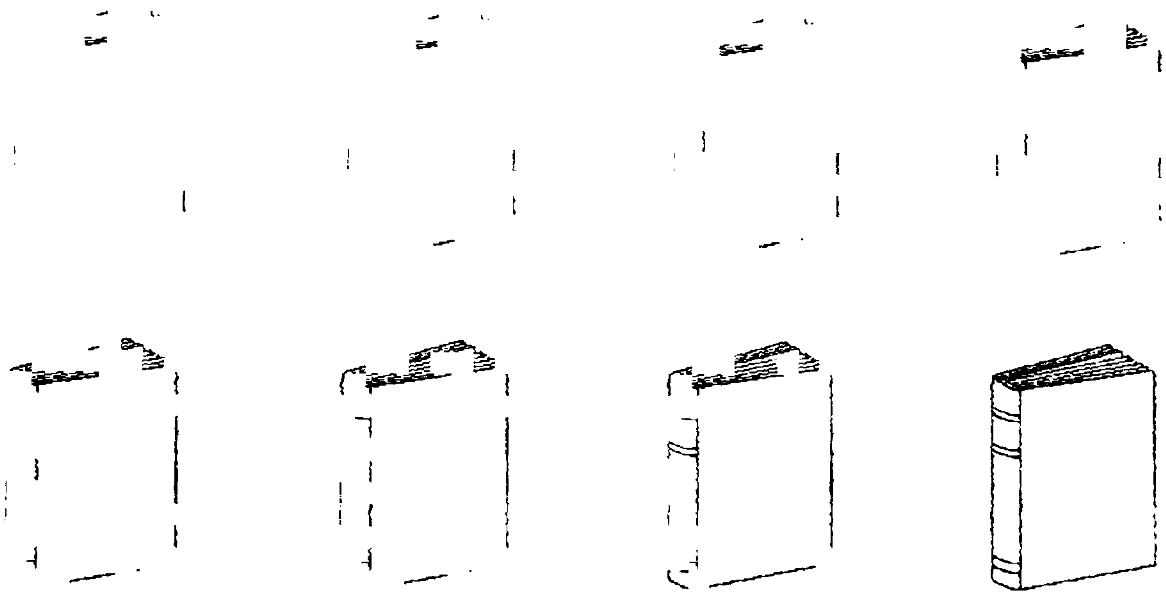


(7) Snowman

(8) Cake



(9) Book



(10) Pear

Appendix N. Sentence Completion Task: Stimuli and scoring examples

	Sentence stems (in order of administration)	Examples of 2-point local completions with 1-point examples underlined	Examples of 0-point global completions
1.	I was given a pen and ... *		
2.	The sea tastes of salt and ...	pepper / vinegar / sugar / sour	water / seaweed / sand / was cold
3.	Hens lay eggs and ...	bacon / chips / milk / noodles / eggs	chicks / have feathers
4.	The woman took the cup and ... *		
5.	You can get burnt by the sun and ...	moon / sea / daughter (son) / sand / stars / rain	fire / hot water / it hurts
6.	You can feed a child bread and ... *		
7.	Little boys grow up to be men and ...	women / lady	girls grow up to be women / adults / granddads
8.	In the sea there are fish and ...	chips	sharks / whales / lots of sea life
9.	In a cave lived a bat and ...	ball	bear / spiders / a caveman
10.	You can go hunting with a knife and...	fork	gun / bow and arrow
	You can swallow apple ... *		
12.	The old shoe-maker mended the shoes and ...	socks / clothes / hats / shirt / laces / slippers	boots / soles / gave them back / cleaned them
13.	The fireman carried the bucket and ...	spade	hose / water / ladder / put out the fire
14.	A vet cares for cats and ... *		
15.	The night was black and ...	white / blue	dark / cold / silver (knight) / had a large sword (knight)

*Control Items

Appendix O. Story Memory Task: Stimuli and scoring system

Test sentences different in surface form to original story are shown in italics

Practice: Football Story

Helen did not play well in the first football game of the season.

After a week of practice, she learned to control the ball better.

Hopefully the coach will let her play the whole game next Saturday.

Football Story Test Sentences

1. Helen did not play well in the first football game of the season.
2. *She learned to control the ball better after a week of practice.*
3. *Hopefully the coach will allow her to play the whole game next Saturday.*

Story 1: Cave Story

Bob and Jane set out to go swimming one day.

They went to a beach they had never been to before.

They swam far out to sea.

Suddenly Bob spotted a big cave in the cliff.

They swam into the cave.

Then they climbed onto a ledge to have a rest.

The swimming had made them very tired and they soon fell fast asleep.

When they woke up the tide had risen.

There was only one way to get out.

They had to swim underwater out of the cave.

They were very lucky to get home safely.

Cave Story Test Sentences

1. Suddenly Bob spotted a big cave in the cliff.
2. *Both of them swam into the cave.*
3. *They found a ledge to have a rest.*
4. *They fell fast asleep because they were so tired from the swimming.*
5. When they woke up the tide had risen.

Story 2: Coin Story

Anne and Tim lived in the same street and were good friends.

One afternoon they were digging in the garden at Tim's house.

After a while Anne dug up a small coin.

Tim found two more coins a little while later.

They dug deeper and deeper looking for more coins.

Altogether they had nine pieces at the end of the afternoon.

They took the coins home to show their parents.

Tim's father took the coins to the museum the next day.

They found out that the coins were very old and very valuable.

Tim and Anne gave them to the museum for everybody to see.

The next day, there was a newspaper article describing their incredible find.

Coin Story Test Sentences

1. *A little while later Tim found two more coins.*
2. *Looking for more coins they dug deeper and deeper.*
3. *At the end of the afternoon they had nine pieces altogether.*
4. They took the coins home to show their parents.
5. Tim's father took the coins to the museum the next day.

Scoring for Gist

	Cave story (maximum =13)
1	set out to go swimming
2	went to a beach / seaside
3	swam far out
4	spotted a cave
5	swam into cave
6	climbed on a ledge / sat on ledge
7	tired from swimming
8	fell asleep
9	woke up
10	tide has risen
11	only one way out
12	swam underwater
13	get home safely

	Coin story: (maximum =11)
1	two friends
2	digging in garden
3	found one coin
4	found two more coins
5	kept on digging
6	found nine coins at the end
7	showed coins to parents / father
8	took coins to the museum
9	found out they were old / valuable
10	gave coins to museum
11	newspaper article about find

Scoring for Local Detail

	Cave story: (maximum =27)
1	Bob
2	Jane
3	set out / got ready / wanted to go
4	swimming / to swim
5	one day
6	went to a beach
7	they hadn't been to before / new beach /
8	they swam
9	far out to sea / a long way / went out
10	Suddenly / Just then

	Coin story: (maximum =34)
1	Anne / Annie
2	Tim
3	lived in the same street / road
4	were good friends
5	One afternoon / One day
6	they were digging / they went digging /
7	in the garden / back garden
8	at Tim's house / at Tim's
9	After a while
10	Anne dug up /Anne found

11 Bob
12 spotted / found / saw
13 a big cave / a cave
14 in the cliff / inside a cliff
15 they swam into the cave / went into it
16 the climbed onto a ledge / went up on a
17 to have a rest / and rested
18 the swimming
19 had made them tired / they were tired
20 they soon fell asleep / went to sleep
21 when they woke up
22 the tide had risen / tide had come up
23 only one way to get out / only way to
24 they had to swim underwater
25 (to get) out of the cave
26 they were very lucky
27 to get home safely

11 a small coin / a coin / an old coin / one
12 Tim found
13 two more coins / two coins / two
14 a little while later / a while later / later on
15 They dug deeper and deeper / they dug
16 looking for more coins / trying to find
17 Altogether
18 they had nine pieces / nine coins
19 at the end of the afternoon / by the end of
20 They took the coins home
21 to show their parents / showed / told their
22 Tim's father / Tim's dad
23 took the coins / took them / took it
24 to the museum
25 the next day
26 They found out / they were told
27 the coins were very old
28 and very valuable
29 Tim and Anne / the children / they
30 gave them to the museum / donated the
31 for everybody to see / so everyone could
32 The next day
33 there was a newspaper article / there was an
34 describing their incredible find / about their

Scoring for Phrase Recall

Cave story: (maximum = 24)

1 Bob and Jane
2 set out to go swimming
3 one day.
4 They went to a beach
5 they had never been to before.
6 They swam far out to sea.
7 Suddenly Bob spotted a
8 a big cave
9 in the cliff.
10 They swam into the cave.
11 Then they climbed
12 onto a ledge
13 to have a rest.
14 The swimming had made them
15 very tired
16 they soon fell fast asleep

Coin story: (maximum = 30)

1 Anne and Tim
2 lived in the same street
3 were good friends.
4 One afternoon
5 they were digging
6 in the garden
7 at Tim's house.
8 After a while
9 Anne dug up
10 a small coin.
11 Tim found two more coins
12 a little while later.
13 They dug deeper and deeper
14 looking for more coins.
15 Altogether they had nine pieces
16 at the end of the afternoon.

17 When they woke up
18 the tide had risen.
19 There was only one way
20 to get out.
21 They had to swim underwater
22 out of the cave
23 They were very lucky
24 to get home safely

17 They took the coins home
18 to show their parents.
19 Tim's father took the coins
20 to the museum
21 the next day.
22 They found out
23 the coins were very old
24 and very valuable.
25 Tim and Anne
26 gave them to the museum
27 for everybody to see.
28 The next day
29 there was a newspaper article
30 describing their incredible find.

Story Memory Task: Inter-rater reliability

The inter-rater reliability of the scoring system used to tally the number of gist elements, local detail, and phrase units recalled was investigated. A second-coder independently rated the Cave story for 30% of the TD sample (61 from 204), 34% of the ASD group (10 from 29), and 37% of the control group (11 from 30). High inter-rater agreement was found across all groups for the three recall elements. Strong correlations between raters were found in the TD sample for gist elements (r_s = .92), local detail units (r_s = .97) and verbatim phrases (r_s = .80, all p < .0005). Similarly, in the ASD group correlation coefficients were high for gist elements (r_s = .96) and local detail units (r_s = .94, all p < .0005), although a lower correlation was found for phrase units recalled (r_s = .67, p = .04). Correlation coefficients for the control group were .74 for gist elements, .99 for local detail units, and .96 for phrase units (all p < .01).

Appendix P. Between-domain Pearson correlations of coherence measures for male and female TD participants

		Low-level auditory	Low-level visuo-spatial		High-level visuo-spatial			High-level verbal		
Males		Chord Sequence	Segment	Integrate	Fragment	Picture Memory	Drawing Style	Sentence	Homo-graphs[a]	Story Memory
Low-level auditory	Segmentation	.04	.07	-.18~	.02	.01	-.09	-.04	.25*	.02
	Chord Sequence		-.24*	-.12	.05	.00	-.07	.05	-.02	-.16
Low-level visuo-spatial	Segmentation			-.06	.04	.16	-.10	.00	.02	.01
	Integration				.24*	-.17~	-.07	.04	.16	-.10
High-level visuo-spatial	Fragmented Pictures					-.20~	-.10	-.09	.10	-.07
	Picture Memory						-.19~	-.19~	-.12	.01
	Drawing Style							.05	-.09	-.15
High-level verbal	Sentence Completion								-.08	-.02
	Homographs[a]									-.15
Females		Chord Sequence	Segment	Integrate	Fragment	Picture Memory	Drawing Style	Sentence	Homo-graphs[a]	Story Memory
Low-level auditory	Segmentation	-.02	.22*	-.09	-.09	-.05	-.01	-.04	.15	-.02
	Chord Sequence		-.23*	.10	.02	-.03	-.07	.11	-.21~	-.03
Low-level visuo-spatial	Segmentation			-.30**	-.14	-.11	.01	-.12	.10	.15
	Integration				.05	.00	.04	.18~	.01	-.23*
High-level visuo-spatial	Fragmented Pictures					-.07	-.04	.03	.03	-.05
	Picture Memory						-.13	-.11	-.06	-.01
	Drawing Style							-.04	.17~	.05
High-level verbal	Sentence Completion								-.10	-.10
	Homographs[a]									-.12

~ $p < .10$, *$p < .05$, **$p < .01$ [a]Extreme outliers removed

Appendix Q. Effect sizes for ASD and Control group comparisons on all coherence indices ranked by magnitude of effect

Task	Coherence Index	Effect size (Cohen's **d**)
Embedded Figures Test	Detection time	0.79
Phoneme Segmentation	Relative effect of position	0.73
Sentence Completion	Local completions error score	0.66
Homograph Reading	Proportion of self-corrections	0.66
Fragmented Pictures	Response time for correct identification	0.54
Navon Similarity Judgment	Total global matches	0.51
Impossible Figures	A' for discrimination im/possible figures	0.39
Picture Memory: Recognition	A' for discrimination of surface form changes	0.30
Story Memory: Recognition	A' for discrimination of surface form changes	0.23
Chord Sequence Task	Number GULR target chords judged correct	0.21
Drawing Style (Rey & Pram)	Coherence composite for copy (order & style)	0.19
Un/segmented Block Design	Relative benefit from segmentation	0.12
Chord Segmentation	Correct detections of target note in a chord	0.06
Chord Segmentation	A' for discrimination of a note within a chord	0.04
Pitch Identification	*Total tone-animal pairs correctly identified*	*-0.10*
Picture Memory: Description	*Global description composite*	*-0.15*
Homograph Reading	*Effect of position (all words, initial response)*	*-0.15*
Phoneme Segmentation	*A' for detection of phoneme in non-words*[a]	*-0.16*
Drawing Style (Rey & Pram)	*Effect of meaning on recall accuracy (Pram minus Rey)*	*-0.35*
Chord Sequence Task	*Number GRLU target chords judged correct*	*-0.54*
Story Memory: Recall	*Proportion of verbatim recall*	*-0.66*

[a]ASD participants with early language delay and their matched controls excluded

Note: negative effect sizes indicate that the direction of the difference in mean performance of the ASD and control groups was opposite to predictions of weak central coherence.